PARAMETRIC MODELLING OF NON-STATIONARY SIGNALS BY LINEAR TIME-DEPENDENT PROCESSES

A Thesis Submitted
in Partial Fulfilment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

By

SUDIPTA MUKHOPADHYAY

to the

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY KANPUR

FEBRUARY 1995

Certificate

It is certified that the work contained in the thesis entitled PARAMETRIC MODELLING OF NON–STATIONARY SIGNALS BY LINEAR TIME–DEPENDENT PROCESSES, by Sudipta Mukhopadhyay has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

Pradip Sircar
Feb. 23, '95

(Dr. P. Sircar)
Assistant Professor
Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur 208 016, INDIA

February 1995

Dedicated

To

My Parents

Remember the words of Christ– "Ask, and it shall be given you; seek, and ye shall find, knock and it shall be opened unto you." These words are literally true –not figures, or fiction.

— *Swami Vivekananda*

Abstract

In this thesis, various aspects of modelling of non–stationary signals are considered. A variety of non–parametric and parametric methods has been introduced in literature to analyse non–stationary signals. The present thesis deals with relating various models to a unified framework and addressing key issues of non–stationary modelling. This, in turn, consolidates the existing techniques and finds new avenues in parametric modelling of non–stationary signals.

The non–stationary signals that we come across in nature vary over time in the following aspects, viz., in frequency content or in energy content or in both. For estimation of evolutionary spectrum by parametric rational transfer function model, so far, attention was focused only on the time variation of AR/ARMA coefficients. The time variation of gain (i.e. the input noise energy) was ignored in this regard. In the present work, it is illustrated that both of those features are important in their own ways. The time variation of AR co-efficients indicates the change in spectral shape with time, and the time–varying gain stores the information about signal intensity. Moreover, the time variation of one feature can not compensate for the time variation of other feature in the estimated evolutionary spectrum. While modelling non–stationary speech signals by time–dependent AR model, the time variation of gain becomes prominent. This feature is really a characteristic of the production mechanism of speech. We present here a non–iterative technique to estimate the gain parameter. Our study reveals that the AR model with time–varying gain is a suitable model for speech phoneme.

Various types of parametric models for non–stationary signals are in use at present. These are the damped sinusoid or complex exponential model, amplitude modulation (AM), frequency modulation (FM) and amplitude/frequency modulation (AFM) models. Damped sinusoids are used for modelling transient signals like nuclear magnetic resonance (NMR) and electromagnetic pulse (EMP) responses. It

is known that the speech signal can be generated by the AFM model. In radar and sonar applications, the FM and AFM signals are encountered.

The main difficulty that we face in employing the above mentioned modulated processes (AM, FM and AFM) to model non–stationary signals is that it leads to a set of non–linear equations, which are to be solved. Our study reveals that all the above mentioned models are only the special cases of the time–dependent ARMA (TARMA) model. Remember here how an undamped sinusoidal signal can be modelled by a constant coefficient AR process and the signal parameters can be estimated through an estimation of the process parameters. It is shown that in the general framework these models keep their own distinctive features, and they can be classified according to the time dependence of position/locus of the pole of the TARMA model. Since various models can be related to the general TARMA model, the equivalence provides a unified approach for synthesis and analysis of the signal. By using a set of basis functions to represent the time variations of the parameters of the TARMA model, one may transform the convoluted non–linear estimation problems to the simple linear ones. This new approach of parametric estimation increases the accuracy of estimation of the base-band signal, compared to what can be achieved by the existing non–parametric methods. The introduction of time–dependent damping factor adds a new dimension in modelling transient data in terms of accuracy/flexibility of the model, as shown in the thesis.

The ECG is an example of a non–stationary signal with its characteristic pseudo–periodicity. The amplitude modulated sinusoidal model is a natural choice for such signal. It is found that when the modulating signal is exponential function of time, it can represent the burst type QRS complex very well. This model not only can synthesise the ECG signal as a whole but also its constituent waves. As the constituent waves of ECG give indications of various activities of the heart, this model may be useful in diagnosis.

The thesis is written to highlight the need of modelling any natural signal according to its own characteristics. The study shows effectiveness and area of application for each modelling technique, and how estimation of parameters for variuos models can be carried out under a unified framework. Finally, the thesis directs attention to identify the topics for future research.

Acknowledgement

First I should pay my tribute to all researchers who have paved the way in non-stationary signal modelling. Among them, my special thanks to Prof. R. M. Loynes for providing the mathematical framework to the rather nebulous concept of evolutionary spectra and to Dr. Y. Grenier for his contribution in transforming the time–dependent linear estimation problem to time–independent linear estimation problem.

I pay my regards to my parents and all my teachers who have contributed a lot to come to my present state. I would thank my guide Dr. Pradip Sircar for all he has done for me. To keep a special note, I must mention about the freedom he has given me in every step of research, ranging from problem selection to thesis writing. This has contributed a lot to build up my personality as a researcher.

I am much thankful to Dr. Debashis Kundu for his patient hearing to my lousy ideas and for helpful suggestions. I am grateful to Prof. G. C. Ray for his immediate help in providing real life data sets, suggestions and encouragement.

My stay in Kanpur has been a pleasant one due to the continuous support from my friends. Due to lack of space I can name only a few of them here. I should thank *TAKURA* (T. K. Roy) and *GOGRA* (G. G. Roy) for their pleasant company and cosy atmosphere in the H–block. Our *Raja* (Anirban Gupta) is a symbol of everlasting energy and goodness. His presence has given me rays of hope whenever I found myself in doldrums. I am thankful to our *His Highness* Amit Neogi for his valuable tips about health care. Due to his sensitive health, he can feel the effect of minute change in food habit and atmosphere and due to his inquisitive nature he has found a lot many hypotheses about the health/mind care. He has kept the information not only with himself, but anyone can benefit from his database about health (and mind) care, absolutely free of charge!! My thanks to *Mighty* Prabal who is always bubbling with energy and has the heart to spend it in proper direction.

To be little more personal, I must thank *Dasji* Biswarup (presently Blue Belt) for giving me company everyday in the afternoon 'Samosha & tea', Tae–Kwan–Do practice and once upon a time in an one year long German Language course. I have enjoyed the taste and power of reasoning of *Puran* (Puranjoy Bhattacharya). I owe a lot to both of them, but for whose help my thesis would not have been as it is today. I want to inform the readers the name of few friends like Joydeb (inventor of Joynet), *Dr.* Kar (Indra), Shibu (the ultimate BOSS of Gnu_plot in IITK!) and loving *Chacha* (Arun Kumar) from whom any one can not only expect but also demand help in need. I regret that I could not exploit them fully for my intellectual growth. Our newly married Dandpat DADA has kindly introduced me to varied things from 'Biriany shop' to 'Biomedical Journals'. I am ever grateful to him. I thank Hall–5 mess for their going out of their way to cater the palate of the residents so that their midday/night research is not hampered by a heavy meal.

At the end, I must thank my caring 'Didi', 'TapanDa' and my cute nieces who have forced me to set an example worth emulating for them.

List of Abbreviation

ACF	auto-correlation function
AFM	amplitude frequency modulation
AIC	Akaike information criterion
AM	amplitude modulation
AR	autoregressive model
ARMA	autoregressive moving average model
ARTG	autoregressive model with time-varying gain
ASPD	averaged spectral distance
DESA	discrete energy seperation algorithm
DFT	discrete Fourier transform
EAM	exponential amplitude modulation
ECG	electro cardio graph
EDC	efficient detection criterion
EMP	electro-magnetic pulse
FIR	finite impulse response
IIR	infinite impulse response
NMR	nuclear magnetic resonence
QSAR	quasi-stationary autoregressive model
RTF	rational transform model
SPDM	spectral distortion measure
STFT	short time Fourier transfer function
TAR	time dependent autoregressive model
TARMA	time dependent autoregressive moving average model
TARTG	time dependent autoregressive model with time dependent gain
TRTF	time dependent rational transfer function
WD	Wigner distribution
WT	Wavelet transfer funtion

Contents

List of Figures

List of Tables

Chapter 1

Problem Statement

1.1 Introduction

In statistical signal processing we often assume the condition of stationarity or wide sense stationarity of the signal under investigation. Under the assumption of stationarity, people were interested in knowing the frequency domain characteristics of the signal. The periodogram along with different types of window are employed to find the power spectral density (PSD) of each signal [1]. After that, the parametric modelling was introduced in determining the PSD. In parametric modelling, the signal is fitted into a known model of unknown parameters, and the analysis is done mainly on the behaviour of the chosen model rather than the signal. The main advantage of parametric modelling are flexibility of analysis and synthesis, and improved resolution in spectrum for short length of data. The improvement or resolution is due to the fact that we are providing more information about the signal. For example, in an autoregressive (AR) model for a finite length of data, one can compute only a finite number of auto-correlation function (ACF) from samples. But with the help of the AR coefficients, we can compute ACFs at any lag from the ACFs computed directly from the samples [1]. In this regard, it should be kept in mind that a bad choice of model can lead to a grossly erroneous result.

It is to be stated, however, that a major weakness of the methods described above is the assumption of stationarity. In fact the real signals are often non–stationarity in nature, and we can name a few of those such as speech, ECG and EEG signals [2]. The first attempt to tackle this problem is to introduce the notion of quasi-stationarity. Here the signal is assumed to be stationary only for a short length of time. This approach helps to justify the validity of the model for the signal. But

as the data length becomes shorter and shorter, the estimate of parameters looses its accuracy. So, one has to compromise between two contradictory requirements to get a workable model. In speech signal analysis, a segment of 25ms is considered to be good for most of the phonemes. But for the quickly varying phonemes such as plosives, the time duration is too large [3]. Hence, one has to compromise with the accuracy of parameter estimate. It seems that the stationary model cannot be expected to give good results when the signal under investigation is non–stationary in nature.

1.2 Literature Survey

1.2.1 Evolutionary Spectrum

For a random process $\{X(t)\}$ whose second order moments are finite and mean is zero, the covariance function R is defined by

$$R(X;\tau,t) = E\{X(t)\overline{X}(\tau)\}, \tag{1.1}$$

where E is the expectation operator and $(\overline{\cdot})$ indicates complex conjugation.

For a stationary process, the spectrum is the Fourier transform of the covariance function. The spectrum gives the information about the distribution of signal energy over the frequency axis. The idea of an *evolutionary spectrum* is just an extrapolation of the concept of spectrum from the stationary case to the non–stationary case. When the process is non–stationary, the covariance function not only depends upon the lag $(\tau - t)$, but also on the instant τ and t. So, it can be said that the evolutionary spectrum will provide information about the distribution of signal energy over the frequency axis at each instant of time. That is in other words, the evolutionary spectrum is a measure of energy density over the time–frequency plane.

In Reference [4], Loynes has given a set of properties, desirable or essential for spectrum Ψ of a process which is not restricted to be stationary. The list is as follows :

A1 : The spectrum Ψ is a real function of time and frequency, and is completely determined by the covariance function.

A2 : Ψ describes the distribution of energy over frequency.

A3 : Ψ transforms reasonably and preferably simply when the process $\{X(t)\}$ is transformed linearly. In particular, a knowledge of the spectrum of X determines the spectrum of the transformed process.

A4 : The relation between the spectrum Ψ and the covariance R is one to one.

A5 : Ψ reduces to the ordinary spectrum, or some simple transform of it, if X is a stationary process.

A6 : If the process is composed of a succession of stationary parts, say $X_1(t)$ if $t \leq 0$ and $X_2(t)$ if $t > 0$, then the spectrum is also composed of the corresponding section of the stationary spectra.

A7 : Ψ is estimatable in principle, probably from one (infinite) length of record.

A8 : Ψ is the Fourier transform, or some related transform, of some apparently meaningful quantity.

Any description of evolutionary spectrum may not satisfy all these properties but **A2** and **A3** seem to form the minimum set which is to be satisfied. Attempts have been made to define the evolutionary spectrum in various ways both parametrically and non–parametrically. Note that our discussion here will consider only the discrete time processes and their evolutionary spectra.

1.2.2 Non–parametric Approach

Similar to the case of stationary signals, for non–stationary signals also a number of non–parametric models were presented initially. The short time Fourier transform (STFT) [5] and Wavelet transform (WT) [6] are examples of this class. Both of these procedures can be looked as transformations from time domain to frequency domain of the windowed signal [7] [8]. One consequence is that the shape of the window has its effect on the energy distribution in time–frequency plane. The other is the fact that the time–frequency resolution is bounded by the uncertainty principle [9].

A familiar non–parametric time–frequency response is Wigner Distribution (WD) which has a large number of desirable properties [10], [11], [12], [13]. Its physical interpretation as energy distribution at any point in the time–frequency plane, however questioned as the WD may assume negative value [10], [14], [15], [16], [17]. Similar objection is raised about its interpretation as energy density at a

point because the measure takes integration over the entire time axis [18]. Moreover, as a quadratic time–frequency distribution it produces interference in case of multicomponent signals [10], [16], [17]. Though considerable amount of research has been done to reduce interference by choosing proper kernel, the attenuation in cross–terms comes only at the cost of broadening of WD signal terms [16], [17].

1.2.3 Parametric Approach

In case of parametric modelling, at first the adaptive models are used for non–stationary signals. Here the first requirement is that the measure of non–stationarity, which may be the rate of variation of the auto-covariance function with time, must be less than the speed of adaptation. Under this assumption, the error in estimation of weight vector has two components. The first one called the weight vector noise is proportional to the speed of adaptation, and the second one called the weight vector lag is inversely proportional to the speed of adaptation [19]. Hence, a compromise has to be made to keep the total error variance minimum. The conflicting requirements of the components of error limits the use of adaptive model in case of highly non–stationary signals.

Next, in Prony's model, the represented signals do not require to be stationary [1]. This model, however, is very sensitive to noise. With the introduction of SVD decomposition [20] and auto-correlation like functions [21], [22], sufficient noise immunity is achieved in the method solved non–iteratively. Some iterative methods are also developed to estimate the parameters in the MLE sense [23]. The technique has got its application in modelling voiced speech signal [24]. This model, however, is useful if the signal is a decaying signal. Otherwise, the model may require large order and the estimate of the model parameters will be difficult [1].

The approach of parametric modelling of non–stationary signals by the use of time–varying rational transfer function (RTF) model [25], [2] significantly does not have the short comings of the other methods, as mentioned earlier. Under certain simplified conditions, the parameters of the time–varying RTF methods can be estimated, and the model has found application in case of natural signals like speech [26] and bat's echo [27].

1.3 Objective of Present Work

A considerable amount of work has been done in parametric modelling of non–stationary signals. This thesis is devoted to the extension, consolidation and integration of various approaches used in this area. The time–dependent RTF model is designed to estimate the evolutionary spectra of non–stationary signals [2]. In the developed method, the main thrust is given to estimate the time–dependent AR coefficients of the fitted system. The issue of time–dependent gain is not properly addressed in this work. In nature, many signals like speech, have time–varying excitations in their production systems [28]. Hence, the use of time–varying gain in estimating the evolutionary spectrum becomes important. In Chapter 2 of this thesis, we have presented a non–iterative technique to estimate the time–varying gain parameter.

It is to be noted that beside the time–dependent RTF model, several non–stationary parametric models are in use. These are Prony's model, the amplitude modulation (AM) model, the frequency modulation (FM) model, and the amplitude frequency modutalion (AFM) model. So, it becomes an interesting issue to know whether these models are equivalent to the time–dependent RTF model, and if so, what are the conditions imposed on the general time–dependent RTF model for the model to be specific. Chapter 3 of this thesis is devoted to this specific issue of equivalence and unification.

The electrocardiogram (ECG) is a real life non–stationary signal. This signal has prominent pseudo-periodicity and burst like QRS complex makes it unfit for forced modelling into a stationary predictive model [29]. The various subsignals of ECG (P, QRS and T) represent different activities of heart [30]. So, from the modelling point of view, it is more meaningful to model these subsignals separately instead of the ECG signal as a whole. It is shown that the amplitude modulation model is a natural choice for this type of signal. Chapter 4 is devoted on the topic of analysis and synthesis of the ECG signal.

To summarise the thesis, it can be stated the various aspects of parametric modelling of non–stationary signals are addressed here. The relationship exhibited between the time–dependent ARMA process and, each of AM, FM, AFM models and Prony's model, provides important insight about non–stationary modelling. The thesis also identifies the problem for further research in this direction.

Chapter 2

Time-Dependent AR Model with Time-Varying Gain

2.1 Introduction

Parametric modelling of non–stationary signal by employing the time–dependent rational transfer function (TRTF) model was proposed by Grenier [2]. The representation of time–varying coefficients by some basis functions gives this approach convenient form, both in terms of estimation and storage. Though the aspect of time variation of gain or changing variance of the noise input is considered at the time of formulation of the problem [2], its implementation together with a workable estimation technique is not found in literature. Even the simplest kind of non–stationary model where the AR coefficients are independent of time and only the gain parameter is varying with time, has not been studied to the best of our knowledge. In case of signals like speech, where the input energy of the production system is time varying [28], the time–varying gain becomes an important model parameter. It is therefore proposed here that the effect of time–varying gain in estimating the evolutionary spectra be studied with the help of both synthetic and natural signals.

It is anticipated that the phonetic segments in speech waveform are characterised by their invariant spectral shapes. As a consequence, gain becomes the only parameter that is time-varying in nature in this case. The suitability of autoregressive model with time–varying gain only (ARTG) is studied for speech phonemes.

2.2 Problem Formulation

Let us consider a non–stationary sequence $x(n)$; $n = 1, 2, \ldots, N$ which is the output of a time–dependent autoregressive process with time–varying gain (TARTG) [2, 31],

$$x(n) + a_1(n)x(n-1) + \cdots + a_p(n)x(n-p) = g(n)w(n) \tag{2.1}$$

for $n = 1, 2, \ldots N$; where $g(n)$ is the time–varying gain, p is the order of the AR process, $w(n)$ is the zero mean white noise input.

Note that the gain $g(n)$ does not have any effect on the locations of the poles. The pole locations are completely determined by the AR coefficients $a_i(n)$. The residues at the poles are decided by both $a_i(n)$ and $g(n)$. Hence, to describe the system completely both $a_i(n)$ and $g(n)$ are to be computed. It should also be noted that the coefficient $a_0(n)$ of $x(n)$ is set to unity in (2.1). This is conventionally done to obtain a non-trivial estimate for $\underline{a} = [a_0(n)\ a_1(n)\ \cdots\ a_p(n)]^T$ while minimising the noise variance of the AR process [32]. To get a non-iterative method for estimation of $g(n)$, let another sequence $y(n)$ be considered such that,

$$y(n) = \frac{x(n)}{g(n)} \tag{2.2}$$

where $g(n) > 0, \forall n$, and $n = 1, 2, \ldots, N$. Combining Equations (2.1) and (2.2) we get,

$$g(n)y(n) + a_1(n)g(n-1)y(n-1) + \cdots + a_p(n)g(n-p)y(n-p) = g(n)w(n) \tag{2.3}$$

We now define a positive nonzero number ϵ such that,

$$\epsilon = \sup_n \left| 1 - \frac{g(n-i)}{g(n)} \right|, \qquad \forall i = 1, \ldots, p. \tag{2.4}$$

In case of real-life sampled signals, the ϵ- factor can be made as small as desired by increasing the sampling frequency f_s and as $f_s \to \infty$, $\epsilon \to 0$. For such a signal, one can get from Equation (2.3),

$$y(n) + a_1(n)y(n-1) + \cdots + a_p(n)y(n-p) \approx w(n) \tag{2.5}$$

which shows that the sequence $y(n)$ is the output of a time–dependent autoregressive (TAR) process with constant gain. The model is of order p, and it is driven by the input sequence $w(n)$. Now the time–varying coefficients can be determined with the technique given in Appendix A [2].

It remains to show how one gets $y(n)$ from $x(n)$, in other words, how the gain factor $g(n)$ is estimated for each n. As it is already assumed that $g(n)$ varies slowly with respect to time, one can estimate it as the local envelope or the short time energy of $x(n)$, when the resulting sequence $y(n)$ is of constant variance. By simulation it is verified that both the interpretations give similar results. Here we have chosen the first approach as it requires less computation. The estimate of $g(n)$ will then be unique up to scaling. Hence, for a symmetric window 'γ' of even length L, $g(n)$ can be estimated as,

$$\hat{g}(n) = \frac{1}{L}\sum_{i=1}^{L}|\gamma(i)x(n+i-\tfrac{L}{2})| \tag{2.6}$$

for $n = \frac{L}{2}, \ldots, (N - \frac{L}{2})$.

The estimated envelope will be approximately unbiased only when the length of the window (L) is less than or equal to the order (p) of the AR model, and $\frac{1}{L}\sum_{i=1}^{L}|\gamma(i)y(n+i-\frac{L}{2})| = 1$. The choice of L depends on the nature of the signal $x(n)$ and that of the sequence $g(n)$. For very small L, the averaging is not complete and the randomness of $x(n)$ is reflected in $\hat{g}(n)$. On the other hand, if L is too large, it fails to follow the local variation of the envelope. As $\hat{g}(n)$ is only unique up to scaling, it may be scaled by a factor $\frac{\max_i |x(i)|}{\max_k g(k)}$ which makes $\hat{g}(n)$ suitable to be called the envelope of $x(n)$.

The non–stationary signal $x(n)$ which is an output of a TARTG is modelled as another non–stationary signal $y(n)$, amplitude modulated by $\hat{g}(n)$ as shown in following Figure 2.1. The signal $y(n)$ is the output of a TAR model with constant gain, and $\hat{g}(n)$ is an envelope function which slowly varies with time.

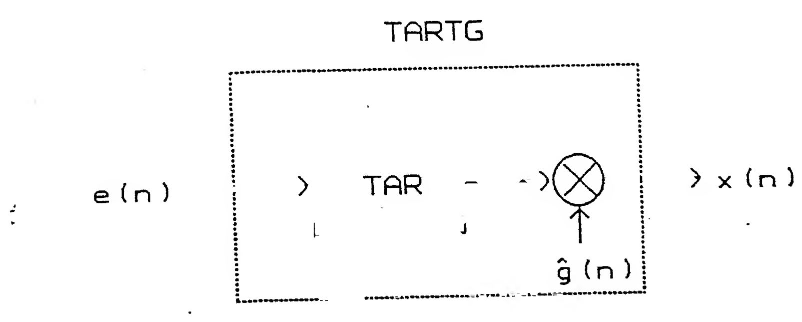


Figure 2.1: The block diagram of time–dependent AR model with time–varying gain (TARTG)

2.2.1 Evolutionary Spectrum

The time–varying transfer function of the system, described by the difference equation (2.1) can be represented as [33],

$$H(n,\omega) = \frac{g(n)}{\sum_{k=1}^{p} a_k(n)exp(-j\omega k)} \tag{2.7}$$

and hence, the evolutionary spectrum of signal $x(n)$ is given by

$$P_{xx}(n,\omega) = |H(n,\omega)|^2\, \sigma_w^2 \tag{2.8}$$

where σ^2 is the variance of the white noise sequence $w(n)$.

It is known that the evolutionary spectrum of $y(n)$ is given by

$$P_{yy}(n,\omega) = \frac{1}{|\sum_{k=1}^{p} a_k(n)exp(-j\omega k)|^2}\sigma_w^2 \tag{2.9}$$

Hence, we find the relationship between the spectra to be

$$P_{xx}(n,\omega) = |g(n)|^2\, P_{yy}(n,\omega) \tag{2.10}$$

2.2.2 Spectral Distortion Measure

A spectral distortion measure (SPDM) is a function of S and $\hat{S}$ which are the true and the estimated spectra respectively of the underlying process. The SPDM assigns a non-negative number $d(S,\hat{S})$ to represent the distortion in using $\hat{S}$ to represent S. As the mean squared error does not appear to be subjectively meaningful in applications like speech coding, the new distortion measures have been introduced.

A distortion measure proposed by Itakura and Saito, and termed as the *error matching measure* is defined as

$$d_{IS}(S,\hat{S}) = \| \frac{S}{\hat{S}} - \ln(\frac{S}{\hat{S}}) - 1 \|_1 \tag{2.11}$$

where S and $\hat{S}$ are true and estimated spectrum of the estimated process and $\|\cdot\|_p$ represents L_p norm.

Gray and Markel proposed a *symmetrised Itakura–Saito distortion measure* termed as the COSH measure of the form [34],

$$d_{COSH}(S,\hat{S}) = \frac{1}{2}[d_{IS}(S,\hat{S}) + d_{IS}(\hat{S},S)] \tag{2.12}$$

For our computation, we have used the *symmetrised Itakura–Saito distortion measure* or the COSH measure. To use it for the evolutionary spectra, the averaged spectral distance (ASPD) is introduced and it is defined as

$$D(S,\hat{S}) = \frac{1}{n}\sum_{i=1}^{n} d_{COSH}(S_i,\hat{S}_i) \tag{2.13}$$

where $\underline{S} = (S_1, S_2, \ldots, S_n)$ and $\underline{\hat{S}} = (\hat{S}_1, \hat{S}_2, \ldots, \hat{S}_n)$ are the true and the estimated evolutionary spectra respectively [34].

2.3 Simulation Results

Example 2.1

Consider an AR model of order $p = 2$,

$$x(n) + \sum_{i=1}^{p} a_i x(n-i) = g(n)w(n)$$

where $w(n)$ = Gaussian white noise with zero mean and unit variance, and

$$g(n) = 1 + 0.5\sin(\frac{\pi n}{1000})$$

The white noise $w(n)$ is generated with the help of a NAG subroutine. The model is started with zero initial condition, and the data of 5000 samples is recorded after 4000 initial samples in order to get rid of the effect of transient. The estimate of $g(n)$, $\hat{g}(n)$ is done with the Hanning window of length $L = 400$ samples. The signal $x(n)$ and its envelope $\hat{g}(n)$ are plotted in Figure 2.2. With the help of $\hat{g}(n)$, the normalised signal $y(n)$ is found, and the AR model is fitted into it. The statistics of coefficients for 100 simulations are given in Table 1, which roughly shows the validity of the model. The evolutionary spectra of the signal and its estimates with the TAR with constant gain and the TARTG are given in Figure 2.3. For the TAR with constant gain a sufficient number (up to 11) of Fourier basis functions are tried, but it is found that no order of basis can produce the time variation of evolutionary spectrum satisfactorily, the time variation being caused by variation in gain. Using the *symmetrised Itakura–Saito spectral distortion measure* (SPDM), we found that the spectral distortion averaged over time (ASPD) of the estimated evolutionary spectra are $7.56X10^{-1}$ and $2.09X10^{-2}$ for the TAR with constant gain and TAR with time–varying gain respectively.

Table 2.1: Bias and variance in estimating the parameters by TARTG model in Example 2.1

Serial i	Parameter a_i	Bias	Variance
1	-0.5	$3.0X10^{-3}$	$2.09X10^{-4}$
2	0.49	$5.38X10^{-3}$	$1.36X10^{-4}$

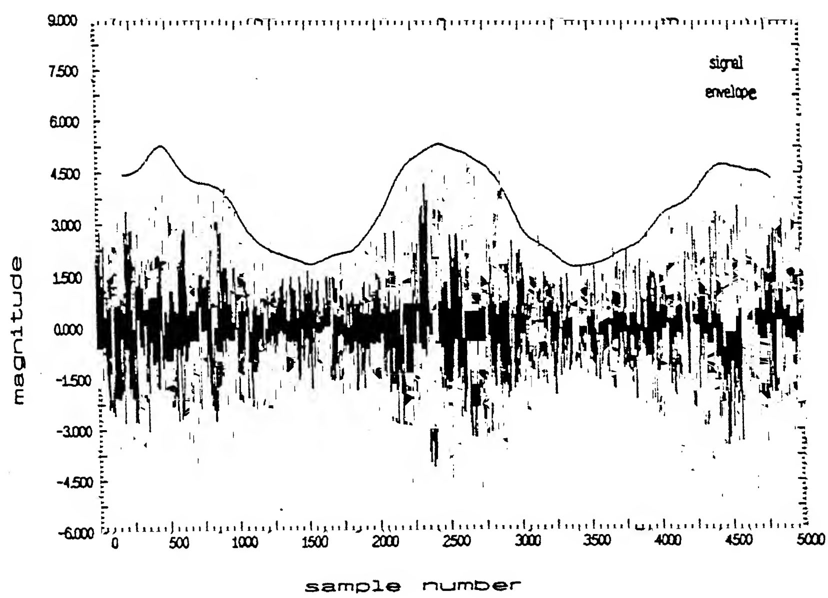


Figure 2.2: The original signal, and estimated envelope of the signal in Example 2.1

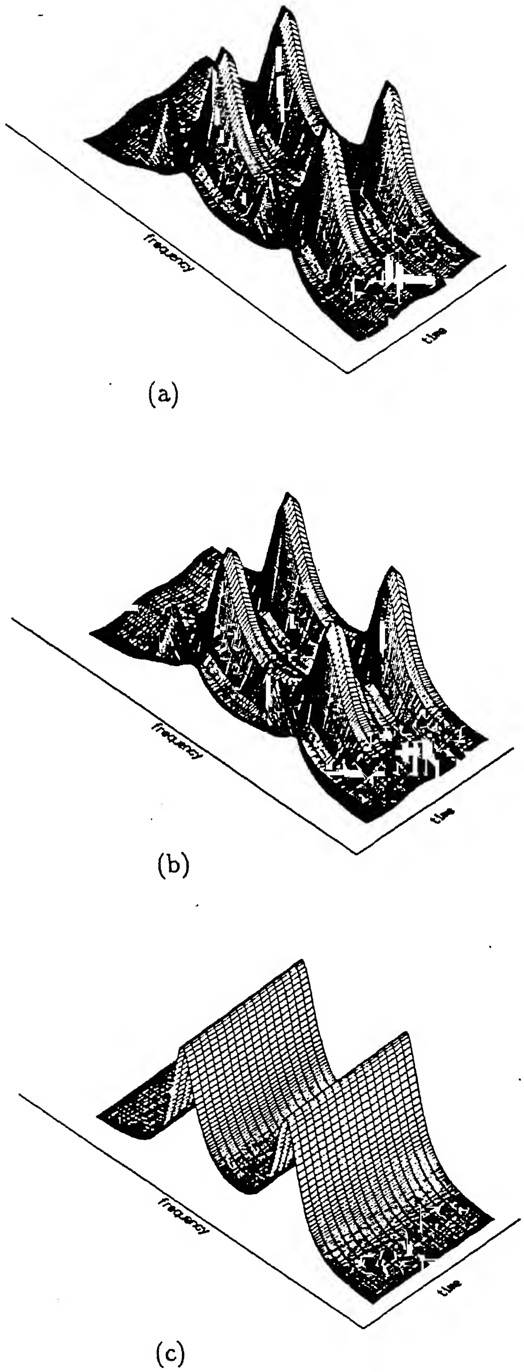


Figure 2.3: The evolutionary spectra of signal in Example 2.1, (a) original; (b) estimated by the TARTG model; (c) estimated by the TAR model.

Example 2.2

In this example, an AR model of order $p = 2$ is considered. The TAR coefficients are chosen as $a_1(n) = -(0.5 - 0.002n)$, and $a_2(n) = 0.49$. Note that the time–varying gain or standard deviation (SD) of white noise is linear in time. The function $g(n)$ is given by

$$g(n) = 1 + \frac{0.25n}{1000}$$

A set of 5000 samples are considered. The estimate of envelope $\hat{g}(n)$ is done with a Hanning window of length $L = 400$ samples. The signal and estimated envelope of the signal are shown in Figure 2.4. The evolutionary spectrum of the signal, and its estimates with the TAR with constant gain and the TAR with time-varying gain are given in Figure 2.5. For the TAR, 2 polynomial basis functions are used, and it is verified that more bases do not effect any appreciable change in the evolutionary spectrum. Using the SPDM, we found the ASPD of the estimated evolutionary spectra are $1.877X10^{-1}$ and $1.82X10^{-2}$ for the TAR with constant gain and the TARTG respectively.

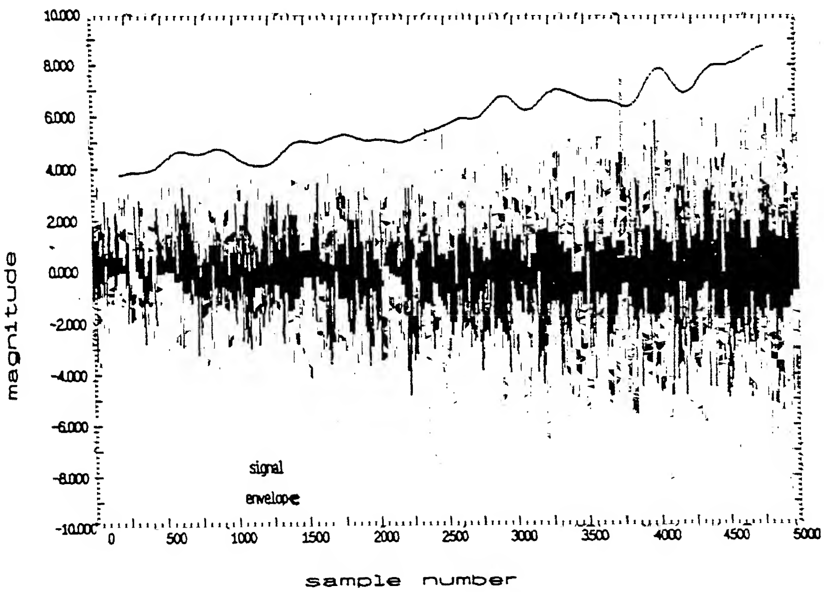


Figure 2.4: The original signal, and estimated envelope of the signal in Example 2.2

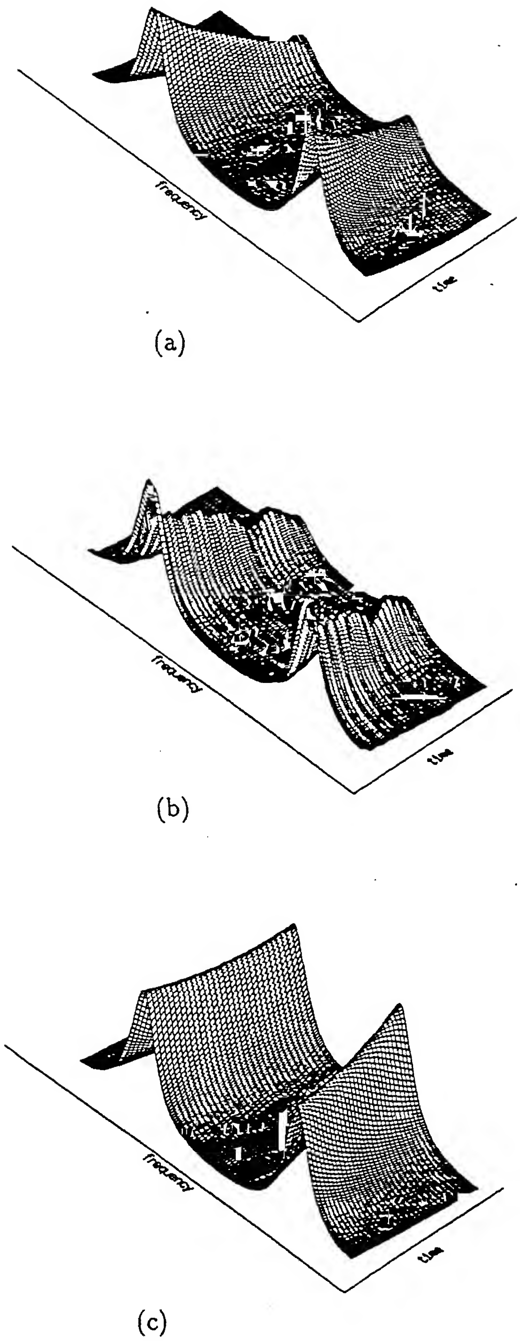


Figure 2.5: The evolutionary spectra of signal in Example 2.2, (a) original; (b) estimated by the TARTG model; (c) estimated by the TAR model.

Example 2.3

The signal considered in this example is a sequence of 6000 samples of the vowel sound /o/ sampled at 16 KHz. The envelope of the signal is estimated by a Hanning window of length $L = 400$. The signal and envelope are plotted in Figure 2.6. Here the original evolutionary spectrum is approximated by the short–time AR spectrum using Hanning window of length 400 i.e. 25 ms. While using the Fourier basis, the TAR model order 10 and number of basis 7 are found to be sufficient for accurate modelling. The approximated original and estimated evolutionary spectra are shown in Figure 2.7.

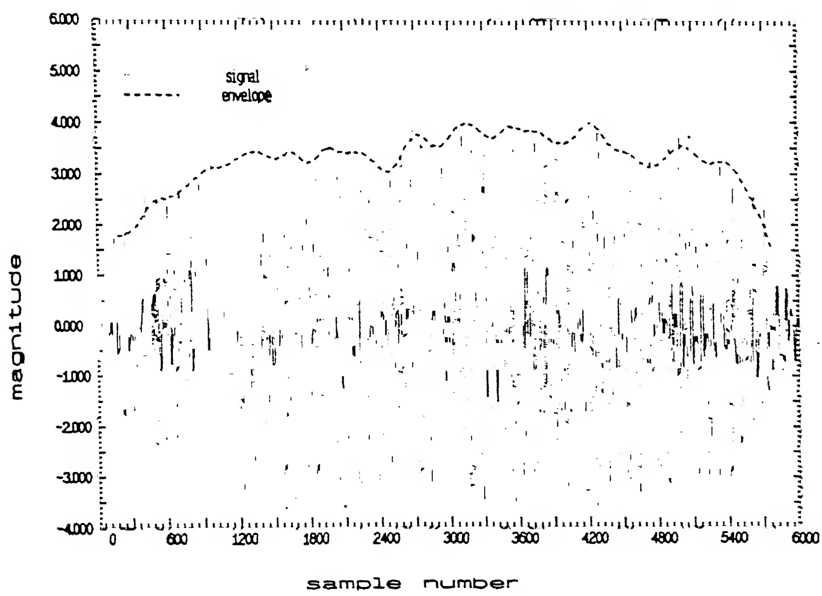


Figure 2.6: The original signal, and estimated envelope of the signal /o/ (vowel) in Example 2.3.

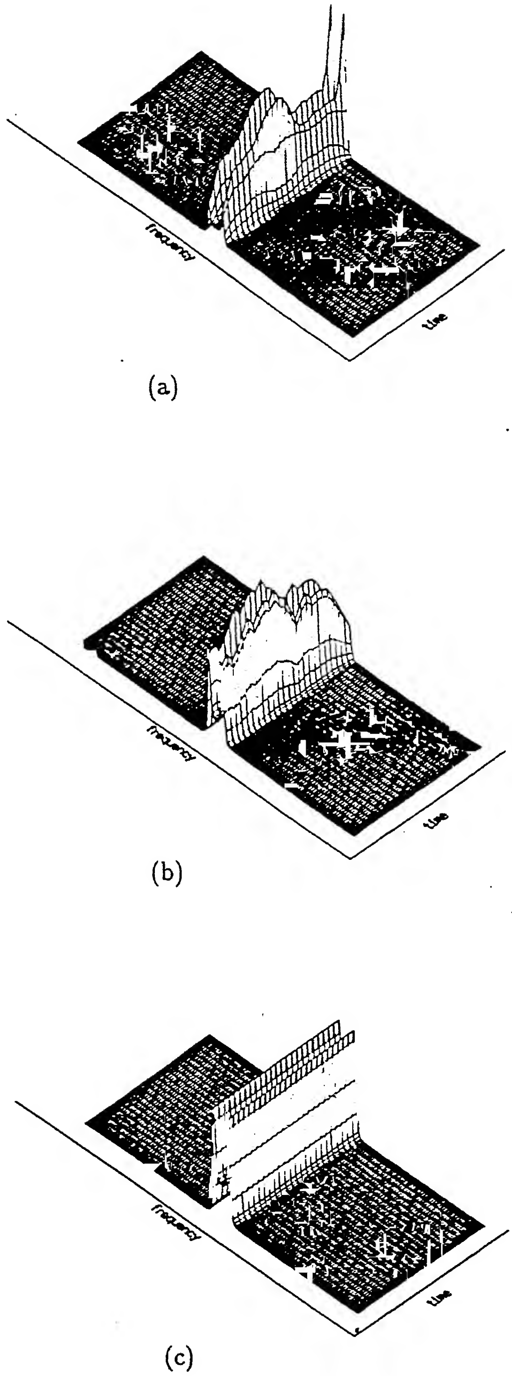


Figure 2.7: The evolutionary spectra of signal /o/ (vowel) in Example 2.3, estimated by (a) the QSAR model; (b) the TARTG model; (c) the TAR model.

Example 2.4

The signal considered in this example is a sequence of 4400 samples of a fricative sound /s/ sampled at 16 KHz. The envelope of the signal is estimated by a Hanning window of length $L = 400$. The signal and envelope are plotted in Figure 2.8. Once again, the original evolutionary spectrum is approximated by the short–time AR spectrum using Hanning window of length 400 i.e. 25 ms. The TAR model order 10 and the number of Fourier basis 7 are found to be sufficient for accurate modelling. The approximated and estimated evolutionary spectra are shown in Figure 2.9.

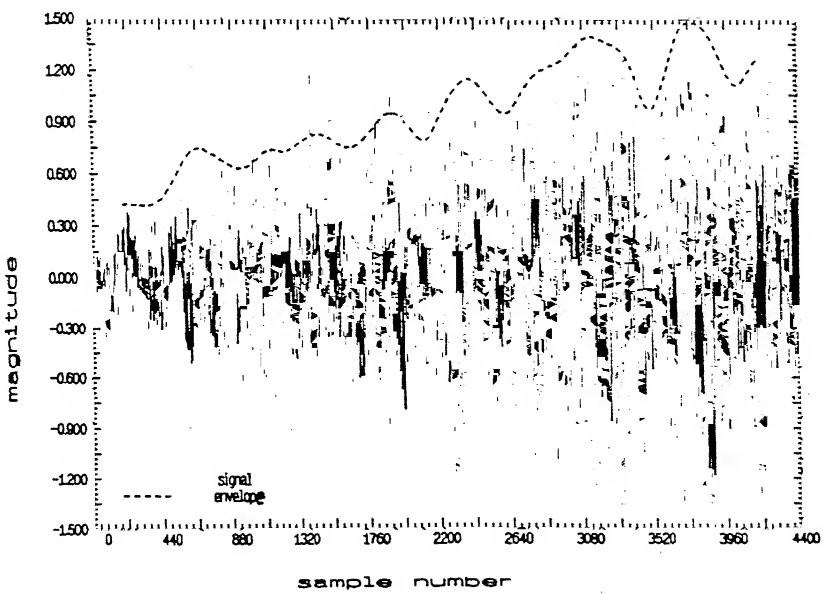


Figure 2.8: The original signal, and estimated envelope of the signal /s/ (fricative) in Example 2.4.

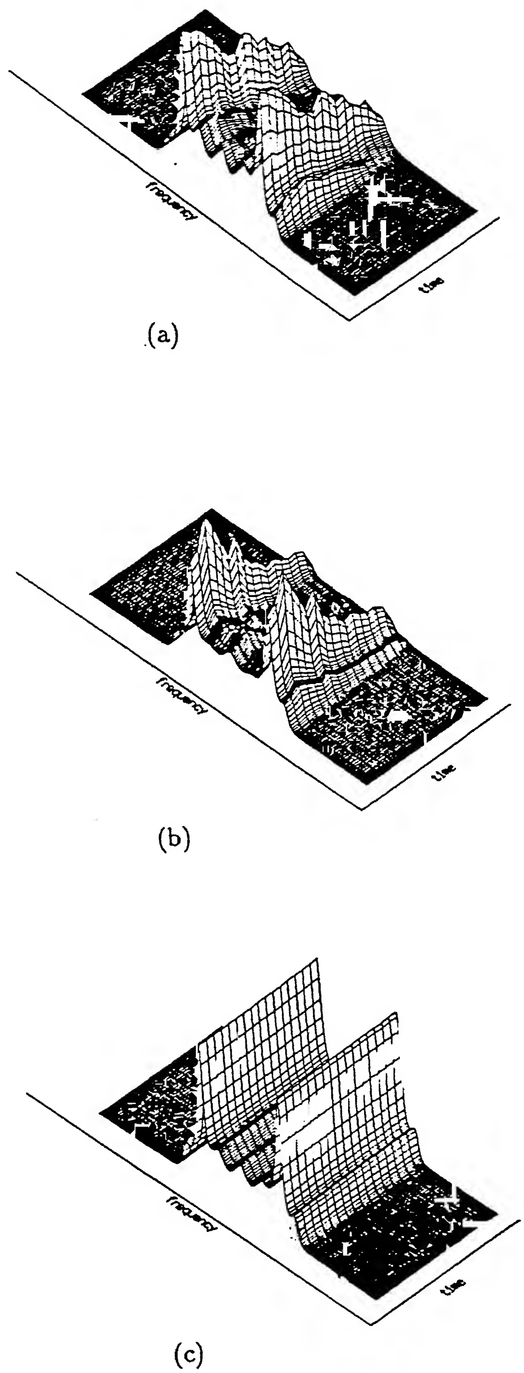


Figure 2.9: The evolutionary spectra of signal /s/ (fricative) in Example 2.4, estimated by (a) the QSAR model; (b) the TARTG model; (c) the TAR model.

2.3.1 Discussion

In all the examples cited in this chapter, the estimated evolutionary spectrum considering the time variation of gain is visually more akin to the original evolutionary spectrum. In Example 2.1, the ASPD of the evolutionary spectrum estimated by the TAR with constant gain is more than 30 times than the ASPD of evolutionary spectrum estimated by the TARTG. In Example 2.2, this factor is 10 for the TAR with constant gain over the TARTG. This clearly shows the superiority of the proposed model in the synthetic examples.

In case of natural signals like speech, the original evolutionary spectra are not known, and hence, quantitative tests are not possible in these cases. Nevertheless, we can see that the estimated evolutionary spectrum by the TARTG is more akin to the approximated evolutionary spectrum by short time AR spectrum than the similar estimate by the TAR with constant gain. This justifies the use of the TARTG in case of natural signals also.

2.4 Modelling of Speech Phonemes by Autoregressive Model with Time–varying Gain

2.4.1 Introduction

Speech is the oldest and most effective mode of communication in human society. In various speech processing problems e.g. coding, recognition, etc., parametric modelling of non–stationary speech signal is an important step. The conventional approach assumes stationarity for the signal over short time-frames, and estimates the set of parameters over each such frame [3]. The usefulness of this approach of modelling, which will be referred to as the quasi–stationary model in this study, becomes limited because the evolutionary nature of signal is not revealed or exploited. This results in an inefficient modelling. The redundancy of information is termed as interframe redundancy and various coding techniques have been devised to exploit it [35].

On the other hand, the time-dependent rational transfer function model developed in References [2, 26], presents a general approach for non–stationary signals.

The method, however, may become less attractive when a large number of basis functions are needed to express the time–variations of the prediction coefficients, and the model order is sufficiently high.

In this thesis, therefore, a middle approach of modelling based on phonetics is proposed for speech signals. Similar concepts are already in use for speech coding, where the need of variable frame length for the quasi– stationary model is advocated [36, 37]. The idea is to relate modelling to linguistic patterns like phonemes, and propose simple non–stationary model valid for each such speech segment.

It is anticipated that phonetic segments in speech waveform are characterised by their invariant spectral shapes. As a consequence, the autoregressive model with time varying gain is proposed in this thesis. The usefulness of the proposed model is demonstrated by extensive simulation with a variety of phonemes.

2.4.2 Model Building

The autoregressive model with time–varying gain (ARTG) is only a special case of TARTG, where all the AR coefficients are time invariant. So, all the Equations (2.1) – (2.10) for TARTG are also valid for ARTG with corresponding change as mentioned above.

Parameterisation of Time-varying gain

When the time–varying gain $g(n)$ is a slowly varying signal, it can be parameterised using a small set of appropriate basis functions. An expansion in terms of cosine functions is found suitable for $g(n)$ and the cosine half series (Appendix B) is truncated where the first turning point of the estimated error variance occurs [38]. In general the change of error variance with the order of expansion is very slow in this study. Hence, an alternative criterion for selecting series–order may be set, which is to have the error variance less than (1/100) fraction of the signal power. If both of these strategies fail, the order is taken as a preset maximum value of 30 in this case.

Order selection of Stochastic process

The most commonly used order–selection criterion of an AR model is the Akaike Information Criterion (AIC). This criterion is known to be an inconsistent one, and it usually overdetermines the model order [39]. Hence, a consistent and more

generalised criterion, called the Efficient Detection Criterion(EDC), is chosen. For AR model, the EDC takes the form [40] :

$$EDC(p) = N \log \sigma_w^2 + C_N p \qquad (2.14)$$

such that

$$\lim_{N\to\infty} (C_N/N) = 0,$$

$$\lim_{N\to\infty} (C_N/\mathrm{loglog}\, N) = \infty,$$

where p, q, σ_w^2 and N are as defined earlier. The penalty function C_N is to be chosen according to the signal under study. In this work, C_N is set to be $C_N = \log N$, for which EDC takes the form of Bayesian Information Criterion [41]. Finally, the order p should be so chosen such that the EDC(p) is minimised.

2.4.3 Results and Discussion

For the purpose of analysis, a total of 65 English phonemes consisting of 8 cardinal vowels and 57 consonants, supplied by *Central Institute of English and Foreign Languages, Hydrabad 500007, India*, is taken into consideration. Each of the cardinal vowels is uttered four times by the same speaker, out of which two sounds are of extended utterance. Each of the consonants is also uttered four times by three different male and one female speakers. The analog data supplied is low–pass filtered with cut off frequency 7.5 KHz and sampled at 16 KHz. As each phoneme gets mixed at both the ends by the following and preceding utterances, a segment of maximum length of uniform characteristic like the average rate of zero–crossing, is chosen by visual inspection of the waveform. To facilitate the comparison of the developed method with the quasi–stationary AR model (QSAR), the length of each segment is restricted to multiples of 25 ms i.e. 400 samples.

With the data sets of different lengths in hand, the first step is to find the time–varying gain $g(n)$. The rectangular window of length 10 ms consisting of 160 samples, is found to be good for this purpose. As the gain is unique up to scaling, it becomes meaningful to call it the envelope of the signal by introducing appropriate scaling factor.The envelope can be parameterised by the cosine half series. The orders of the cosine series are calculated for each phoneme and the average is computed. The average number (= 18) of basis functions needed is highest for vowels, and it is almost double than the average number for consonants.

The average series–order over all speakers and all phonemes is 10. In Figure 2.10, it is shown how the choice of the number of basis functions affects the closeness of approximation of the envelope. The highest order of 30 provides the best result, whereas the order 10 produces too smooth an envelope.

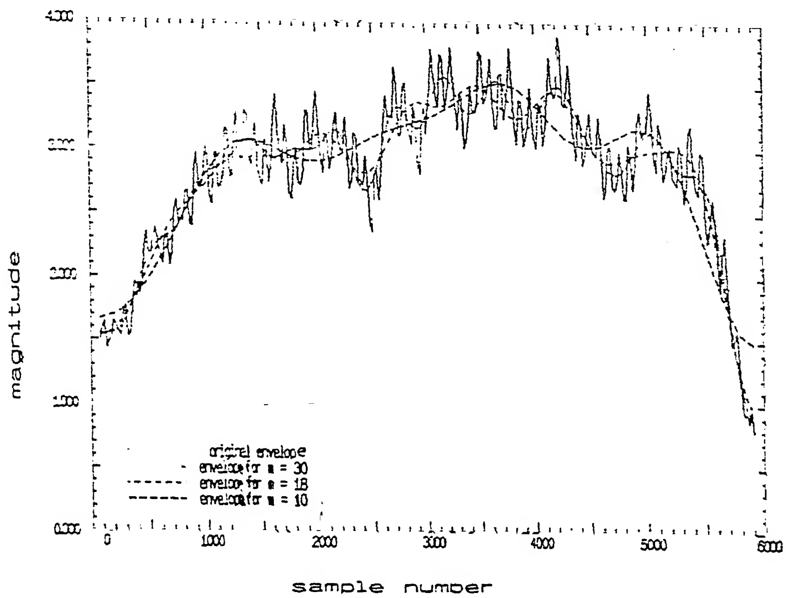


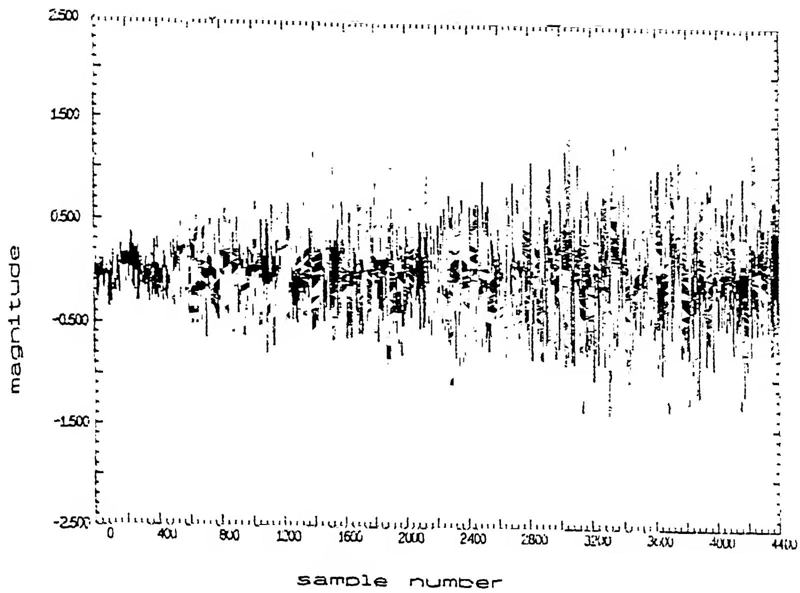
Figure 2.10: The estimated envelope of the signal /o/, and its approximations for 30,18 and 10 basis functions of cosine half series.

Estimating first the gain $g(n)$, the normalised sequence $y(n)$ can be determined and then the corresponding AR parameters can be estimated. By using the EDC, the optimal order of the process is determined and the corresponding SNR is calculated. The SNR is defined as

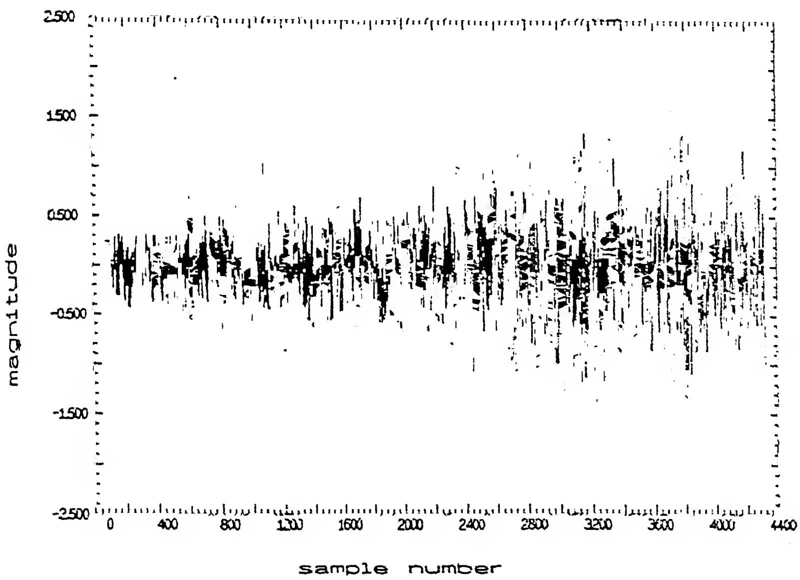
$$\text{SNR} = 10 \log \frac{\text{signal power}}{\text{prediction error power}}$$

For the fricative /s/, the original signal and the signal regenerated by the ARTG are shown in Figure 2.11. For the ARTG, the average of optimal AR orders for all the phonemes is found to be 8, and the average SNR is calculated to be 26 dB. For the vowel sounds, the average SNR is 36 dB, and it is 24 dB for consonants.

The merit of the ARTG model may be judged in relation to the other existing methods. For the quasi–stationary AR model (QSAR) the stationarity is assumed over a period of 25 ms i.e. for 400 samples [3]. For the non–stationary AR model



(a)



(b)

Figure 2.11: The signal /s/, (a) the original, and (b) the synthesised one by ARTG.

(NSAR), the prediction coefficients are all functions of time [2, 26]. The order is chosen to be 10 in each case. In case of the NSAR model, 11 basis functions for each coefficients are employed, hence a total of $(10 \times 11 + 1 = 111)$ parameters are needed for each phoneme. For the QSAR model, the number of parameters will be $(10 + 1 = 11)$ per block, where a block consists of 400 samples over a duration of 25 ms. For the ARTG model, the number of basis functions is chosen to be 15 instead of the average basis–order of 10. This is necessary to get reasonably good approximation in case of vowels. Hence for the ARTG, a total of $(10+15+1 = 26)$ parameters per phoneme is required. The achievable SNRs for all the models are tabulated in Table 2.2, whereas the number of parameters required per phoneme in each case is given in Table 2.3. Table 2.2 shows that the ARTG performs almost same as the NSAR model (0.5 dB less) and little inferior ($\sim$ 2 dB) to the QSAR model. It can be shown from Table 2.3 that the QSAR model usually needs the maximum and the ARTG needs the minimum number of parameters to represent a phoneme. If the number of parameters per phoneme required by the ARTG is taken to be 1 p.u., then the QSAR and NSAR models need 4.95 p.u. and 4.27 p.u. respectively.

2.5 Remarks

The study reveals that if a non–stationary signal has a variation in energy, then the time–varying gain must be incorporated in modelling the signal. In this way, a faithful estimate of the evolutionary spectrum of the signal can be obtained. It is clear that time variation of the AR process cannot represent the effect of variation in gain. Though the proposed modification costs a little more computation, it makes the model more meaningful. In many processes of signal generation, the input signal energy is known to vary with time e.g. in speech signal production [28]. Recent study shows that the proper understanding of the evolutionary spectrum helps to enhance the performance of coders for speech and image signals to a great extent [42].

It is revealed in the study that the ARTG model is more suitable for speech signal than the QSAR model or TAR model with regard to the parsimony of parameters.The SNR for the ARTG is similar to that of the TAR model, and marginally inferior to the SNR of the QSAR model. The evolutionary spectrum computed by the ARTG is similar to that by the QSAR model, suggesting that the ARTG

Table 2.2: SNR in modelling phonemes by QSAR, TAR and ARTG models

Phoneme		SNR $\mu \pm \sigma$ dB		
Name	Number	QSAR	TAR	ARTG
Plosive	13	28.78 ± 9.22	27.22 ± 9.2	26.17 ± 8.95
Nasal	7	30.36 ± 8.53	28.35 ± 8.53	28.30 ± 8.83
Trill	2	24.51 ± 7.3	23.35 ± 7.88	21.49 ± 5.4
Tap or Flap	2	27.97 ± 8.49	27.01 ± 8.76	26.73 ± 8.62
Fricative	22	23.28 ± 8.49	21.96 ± 8.37	21.22 ± 8.17
Lateral Fricative	2	23.26 ± 7.19	22.3 ± 7.48	20.72 ± 6.79
Approximant	5	29.13 ± 9.52	27.22 ± 9.45	27.52 ± 9.73
Lateral Approximant	4	30.15 ± 8.89	28.31 ± 8.93	27.43 ± 8.91
Vowel	8	37.8 ± 5.06	35.12 ± 4.88	35.84 ± 5.44
Average		27.98	26.32	25.8

Table 2.3: Number of parameters required by QSAR, TAR and ARTG models for phonemes

Phoneme		Number of Parameters per Phoneme		
Name	Duration $\mu \pm \sigma$ ms	QSAR	TAR	ARTG
Plosive	246.9 ± 147.43	108.64	111	26
Nasal	243.83 ± 88.73	107.28	111	26
Trill	268.75 ± 124.84	118.25	111	26
Tap or Flap	184.38 ± 128.05	81.13	111	26
Fricative	264.77 ± 150.64	116.5	111	26
Lateral Fricative	218.75 ± 93.33	96.25	111	26
Approximant	277.5 ± 143.81	122.1	111	26
Lateral Approximant	256.25 ± 142.11	112.75	111	26
Vowel	564.84 ± 246.6	248.53	111	26
Average	292.56	128.72	111	26

is the best choice to model phonemes. The entire study suggests that within a phoneme the change in spectral wave shape is small whereas the short–time energy can have a large variation. Figures 2.7 and 2.9 also support the above-mentioned fact. Moreover, the time–varying gain or the envelop estimated as a part of the model identification, can be of use in the recognition of consonants [43, 44]. Therefore, the ARTG seems to be quite promising in modelling phonemes, which can in turn be used directly in speech synthesis. Moreover, the ARTG model together with a phoneme boundary detection algorithm can be used for coding and recognition of the continuous speech.

Chapter 3

Modelling Non–Stationary Signals : A Unified Approach

3.1 Introduction

Under the general framework of time–dependent ARMA (TARMA) model, any signal can be represented as the output of a system with time–varying poles. It is revealed in this thesis that an interesting classification of the non–stationary signals is possible through functional specifications of the movements of time–varying poles. In this regard, this study presents a unified approach to encompass various techniques of parametric modelling of non–stationary signals viz. Prony's model, the amplitude modulation (AM), frequency modulation (FM) and, amplitude and frequency modulation (AFM) models [21, 45, 46, 47, 48].

In my study, it is found that the damped sinusoid, AM, FM and AFM sinusoids can be generated from the TARMA model by using proper restrictions on the model parameters. This finding provides a general procedure for estimating the parameters of various non–stationary models using a set of basis functions. In this way, the estimation of parameters can be obtained by solving a set of linear equations as shown in Appendix A [2]. It is demonstrated that the unified approach of signal modelling and parameter estimation results in greater modelling flexibility (choice of more suitable model) and/or accuracy. The better accuracy is achieved because the method presented here transforms the nonparametric/nonlinear estimation problems into the respective parametric/linear estimation problems.

This thesis reveals that the impulse driven time–dependent all pole models viz. the AM, FM, AFM and damped sinusoidal signal models can be visualised as some

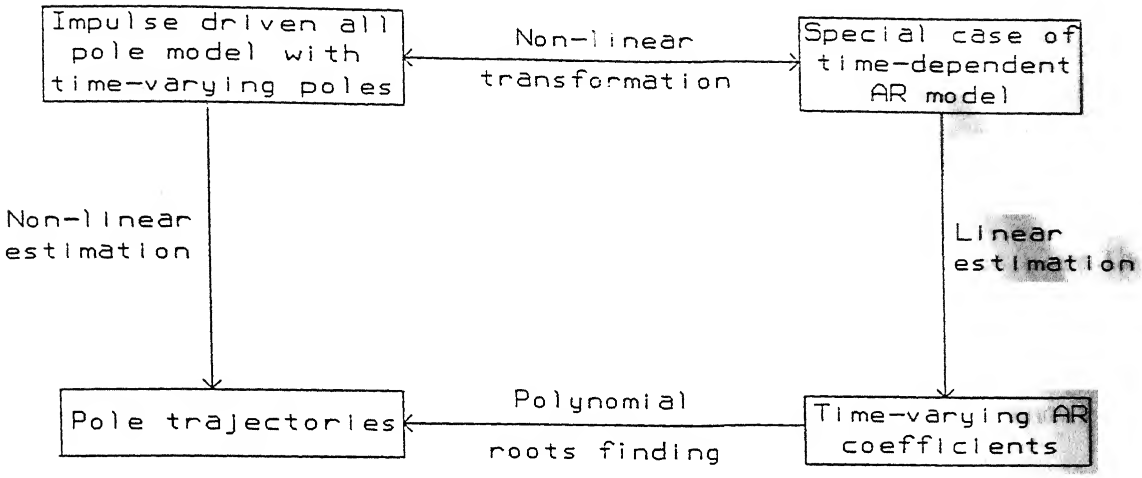


Figure 3.1: Alternative non–stationary models

special cases of the time–dependent AR (TAR) model. This fact is explained in the following chapter and the interrelation among various models is shown in Figure 3.1.

The trajectory of pole of the impulse driven system keeps information of the modulating signal in case of the modulating signal model. For the damped sinusoidal model, the pole trajectory provides information about the rate of decay of the signal. Note that the direct estimation of poles gives rise to a non–linear estimation problem even in case of stationary signals. The situation becomes much more complicated with non–stationary signals.

It is to be pointed out that although the TAR coefficients contain the same information about the modelled signal, the coefficients do not provide any characterisation for the signals. But the TAR coefficients can be estimated by solving a set of linear equations, when the coefficients are expressed in terms of some suitable basis functions. This motivates us to estimate the TAR coefficients and then, to find the pole trajectories by solving the roots of the polynomial. In this way, we get a meaningful characterisation of the signals through a relatively straightforward estimation technique.

3.2 Model Building

3.2.1 Problem Formulation

The input/output relation of the process is given by [31, 49],

$$x(n) + a_1(n)x(n-1) + \cdots + a_r(n)x(n-r) = b_0(n)w(n) + \cdots + b_q(n)w(n-q) \quad (3.1)$$

where $a_k(n) : k = 1, \ldots, r$ are the AR parameters, $b_l(n) : l = 0, 1, \ldots, q$ are the MA parameters, $w(n)$ is the zero mean white noise input, and $x(n)$ is the output sequence of the process. The sequence $x(n)$ is assumed to be non–stationary.

The generalised transfer function $H(n,z)$ corresponding to the difference equation (3.1) can be expressed as [49],

$$H(n,z) \cong \frac{\sum_{l=0}^{q} b_l(n) z^{-l}}{\sum_{k=0}^{r} a_k z^{-k}} \quad (3.2)$$

where it is assumed that for all n,

$$H(n,z) \simeq H(n-k,z), \; for \; k = 1, 2, \ldots, r \quad (3.3)$$

The condition imposed by (3.3) is generally valid for slowly varying system, and it is usually less restrictive than the assumption of stationarity for $x(n)$ over a data-window. Note that, the model order is usually much small compared to the window size in case of the quasi–stationary AR model or the short time Fourier transform.

The time–varying rational transfer function in (3.2) can be rewritten as [49],

$$H(n,z) = \frac{g(n) \prod_{l=1}^{q} (1 - z_l(n) z^{-1})}{\prod_{k=1}^{r} (1 - p_k(n) z^{-1})} \quad (3.4)$$

or,

$$H(n,z) = \sum_{k=1}^{r} \frac{c_k(n)}{1 - p_k(n) z^{-1}} \quad (3.5)$$

where $g(n)$ is the time–varying gain of the generalised transfer function, $p_k(n) : k = 1, \ldots, r$ are the time–varying poles, $z_l(n) : l = 0, 1, \ldots, q$ are the time–varying zeros, $c_k(n)$ are the time–varying residues at the poles $p_k(n)$, and $r > q$ for realisability.

While writing (3.5), it is assumed that the rational transfer function has no repeated pole and the order of zero is less than the order of pole. Under these conditions, a rational transfer function can be written as a weighted sum of single-pole systems, where weights may vary with time.

It will, therefore, be instructive to study the first order time–varying AR model and understand the multiple-pole system as a combination of single-pole systems. Let a first order AR process be given by

$$x(n) = \begin{cases} ax(n-1) + w(n) & \text{if } n < n_0 \\ ad(n)x(n-1) + w(n) & \text{if } n \geq n_0 \end{cases} \tag{3.6}$$

where $w(n)$ is the white noise input with zero mean and variance σ_w^2.

Note that initially the coefficient 'a' is set at $a = (1.0 - \epsilon)\exp(j\theta)$ with positive ϵ and the variance of the white noise is fixed at σ_w^2. The values of ϵ and σ_w^2 are made sufficiently small so that $x(n)$ becomes a sinusoidal signal before n_0. This feature is explained in Appendix C [1]. At the time $n = n_0$ with n_0 large enough to assure complete decay of transients, the time–varying factor $d(n)$ is introduced and its effect on $x(n)$ is investigated.

In (3.6) as $|a| \to 1$ and $\sigma_w^2 \to 0$, the time series $x(n)$ grows to a sinusoid with amplitude much bigger than σ_w [1]. So, for $n \geq n_0$, the contribution of $w(n)$ in $x(n)$ is negligible compared to the contribution of $x(n-1)$. Then the skeleton structure [50, Chapter 4] of (3.6) becomes

$$x(n) = ad(n)x(n-1), \; n > n_0 \tag{3.7}$$

At $n = n_0 + m$ and with repeated substitution leads to

$$\begin{aligned} x(n_0 + m) &= a^m \prod_{l=n_0+1}^{n_0+m} d(l)x(n_0) && (3.8) \\ &= a^m D(m)x(n_0) && (3.9) \end{aligned}$$

where $D(m) = \prod_{l=n_0+1}^{n_0+m} d(l)$.

3.2.2 Classification of models

It is shown in this section that the use of some mathematical functions for $d(n)$ may lead to a few well known models for non–stationary signals. These models are,

- Damped sinusoidal model,
- Amplitude modulated sinusoidal model,
- Frequency modulated sinusoidal model,
- Amplitude and Frequency modulated sinusoidal model.

In the following subsections, these models are investigated in detail.

Damped Sinusoidal Model

In this case, the time–varying coefficient is set to be

$$d(n) = b^{-(n-n_0)}, \; b > 1, \; n > n_0 \tag{3.10}$$

Substituting $d(n)$ in (3.9) and after simplification we get

$$D(m) = b^{-\frac{m(m+1)}{2}}$$

and

$$x(n_0 + m) \simeq b^{-\frac{m(m+1)}{2}} x(n_0) \exp{(j\theta m)} \tag{3.11}$$

where $a \approx 1.0 \exp{(j\theta)}$.

So, for $d(n)$ as defined in (3.10), the process decays rapidly as indicated by (3.11), and after sometime, the output becomes same as the input noise sequence.

Note here that as a special case, if we set

$$d(n) = b^{-1}, \; b > 1 \; \& \; n > n_0 \tag{3.12}$$

from (3.9) we get

$$D(m) = b^{-m}$$

and

$$x(n_0 + m) \simeq b^{-m} x(n_0) \exp{(j\theta m)} \tag{3.13}$$

where $a \approx 1.0 \exp{(j\theta)}$. The system loses its energy exponentially and settles to a new AR model with pole at 'ab'.

The difference between $d(n)$ in (3.10) and (3.12) is that the damping factor given by $\ln |d(n)|$ decreases linearly in the first case and it is constant in the second case. This not only results in rapid decay of the signal in the first case, the rate of decay also increases with time.

Amplitude Modulated Sinusoidal Model

In this case, the time–varying coefficient is substituted as

$$d(n) = \exp{(b \sin \alpha(n - n_0))}, \; n > n_0 \tag{3.14}$$

which after substitution in (3.9) and with subsequent simplification leads to

$$\begin{aligned} D(m) &= \prod_{l=n_0+1}^{n_0+m} d(l) \\ &\simeq \exp\left(-\frac{b}{\alpha}(\cos \alpha m - 1)\right) \end{aligned} \tag{3.15}$$

The approximation of the above equation holds true for small α (Appendix D). Rewritten,

$$x(n_0+m) \simeq 1.0 \exp(j\theta m) \exp\left(\frac{b}{\alpha}(1-\cos\alpha m)\right) x(n_0) \tag{3.16}$$

where $a \approx 1.0 \exp(j\theta)$ is substituted.
Taking real part of the above equation, one gets

$$real[x(n_0+m)] \simeq A \exp\left(\frac{b}{\alpha}(1-\cos\alpha m)\right) \cos(\theta m+\phi) \tag{3.17}$$

where $x(n_0) = A\exp(j\phi)$. Note that (3.17) gives the output as an amplitude modulated signal with modulating factor given by $\exp(-\frac{b}{\alpha}\cos\alpha m)$.

For any positive discrete–time signal $u(n)$, we can modulate with the factor

$$d(n) = \frac{u(n)}{u(n-1)} \tag{3.18}$$

and, we get the desired output in discrete time domain as given by

$$x(n_0+m) = u(m)\exp(j\theta m)x(n_0)/u(n_0) \tag{3.19}$$

The time–varying part of the pole $d(n)$ is real in the case of amplitude modulation. This results in time variation of the magnitude of the pole, keeping its angle constant. In other words, for $a \approx 1.0 \exp(j\theta)$, the pole will have radial movement alternating through the point $a \approx 1.0\exp(j\theta)$, in this case.

Frequency Modulated Sinusoidal Model

In this case we assign the factor $d(n)$ to be

$$d(n) = \exp(jb\sin\alpha(n-n_0)), \quad n > n_0 \tag{3.20}$$

which after substitution in (3.9) and with subsequent simplification leads to

$$\begin{aligned} D(m) &= \prod_{l=n_0+1}^{n_0+m} d(l) \\ &\simeq \exp\left(-j\frac{b}{\alpha}(\cos\alpha m-1)\right) \end{aligned} \tag{3.21}$$

The approximation of the above equation holds true for small α (Appendix D). From (3.21) we obtain

$$x(n_0+m) \simeq 1.0\exp(j\theta m)\exp\left(j\frac{b}{\alpha}(1-\cos\alpha m)\right) x(n_0) \tag{3.22}$$

where $a \approx 1.0 \exp(j\theta)$ is substituted as before.
The real signal is given by

$$real[x(n_0 + m)] \simeq A \cos\left(\theta m - \frac{b}{\alpha} \cos \alpha m + \phi + \frac{b}{\alpha}\right) \quad (3.23)$$

where $x(n_0) = A \exp(j\phi)$. It is clear that (3.23) gives the singletone frequency modulated output of the process.
Therefore, in order to modulate with a discrete-time signal $u(n)$, one has to use

$$d(n) = \exp(jKu(n)) \quad (3.24)$$

for FM, which leads to the FM signal

$$x(n_0 + m) = A \exp\left(j(\theta m + K \sum_{l=n_0+1}^{n_0+m} u(l) + \phi)\right),$$

and

$$d(n) = \exp(jK\dot{u}(n)) \quad (3.25)$$

for PM, where $\dot{u}(n)$ is the first order backward difference of $u(n)$. The corresponding PM signal is

$$x(n_0 + m) = A \exp(j(\theta m + Ku(n) + \phi - Ku(n_0)).$$

The time–varying part of the pole $d(n)$ is an imaginary quantity in the case of frequency modulation. This means that $d(n)$ will cause only the angular movement of the poles. In other words, for $a \approx 1.0 \exp(j\theta)$, the poles will have circumferential movement alternating around the point $a \approx 1.0 \exp(j\theta)$.

Amplitude and Frequency Modulated Sinusoidal Model

In this case, we apply both the earlier mentioned schemes combinedly. The proposed $d(n)$ will be of the form,

$$d(n) = \frac{u_1(n)}{u_1(n-1)} \exp(ju_2(n)) \quad (3.26)$$

where the positive discrete-time signal $u_1(n)$ is amplitude modulating signal and $u_2(n)$ is the frequency modulating signal.

3.2.3 Estimation of Parameters

As a complex sinusoid can be represented by a first order AR model, its real part will require a second order AR model. Let us have the characteristic equation corresponding to a first order time-dependent complex AR model (3.6) for $n \geq 0$;

$$1 - a(n)z^{-1} = 0 \tag{3.27}$$

where $a(n) = ad(n)$ and z^{-1} is the delay operator [25, 51] .
The characteristic equation of the corresponding real second order AR model is then

$$\begin{aligned} \left(1 - a(n)z^{-1}\right)\left(1 - a^*(n)z^{-1}\right) &= 0 \qquad (3.28)\\ \text{or,} \quad 1 - \left(a(n) + a^*(n)\right)z^{-1} + a(n)z^{-1}a^*(n)z^{-1} &= 0 \\ \text{or,} \quad 1 - \left(a(n) + a^*(n)\right)z^{-1} + a(n)a^*(n-1)z^{-2} &= 0 \\ \text{or,} \quad 1 - \left(a(n) + a^*(n)\right)z^{-1} + |a(n)|^2 z^{-2} &\simeq 0 \end{aligned}$$

where $a^*(n)$ is the complex conjugate of $a(n)$ and the system is assumed to be slowly varying in time as in (3.3). Let the second order time-dependent AR model to represent a real sinusoid be,

$$x(n) + a_1(n)x(n-1) + a_2(n)x(n-2) = w(n) \tag{3.29}$$

where $w(n)$ is the prediction error which takes into account all the sources of errors viz. the noise in the data, error in approximating the time-dependent prediction coefficients, etc. The sequence $w(n)$ is assumed to be of zero mean and variance σ_w^2. Comparing (3.28) with (3.29) we get

$$\begin{aligned} a_1(n) &= -\left(a(n) + a^*(n)\right) \qquad (3.30)\\ &= -2|a(n)|\cos\left(\arg \mathrm{a}(\mathrm{n})\right) \\ &= -2|d(n)|\cos\left(\theta + \arg \mathrm{d}(\mathrm{n})\right) \end{aligned}$$

$$\begin{aligned} a_2(n) &\simeq |a(n)|^2 \qquad (3.31)\\ &\simeq |d(n)|^2 \end{aligned}$$

where $a \approx 1.0 \exp(j\theta)$ and $d(n) = |d(n)| \exp(j\arg \mathrm{d}(\mathrm{n}))$.
From (3.30) and (3.31) we get,

$$|d(n)| \simeq \sqrt{a_2(n)} \tag{3.32}$$

$$\arg d(n) \simeq \arccos\left(\frac{-a_1(n)}{2\sqrt{a_2(n)}}\right) - \theta \tag{3.33}$$

From (3.18) the estimate of the modulating signal for amplitude modulation can be expressed as

$$\hat{u}_1(n_0+m) = \prod_{l=n_0+1}^{n_0+m} |d(l)| u_1(n_0) \tag{3.34}$$

and from (3.24) the estimate of the modulating signal for frequency modulation can be given as

$$\hat{u}_2(n) = \arg d(n) \tag{3.35}$$

Both these estimates give only the scaled and shifted version of the modulating signals. Note that $|d(n)|$ represents the radial movement of the system pole and the quantity keeps information about the amplitude modulating signal. Whereas, $\arg d(n)$ represents the circumferential movement of the pole and it keeps information of the frequency modulating signals.

Since we can only get some scaled and shifted estimates of the modulating signals, as given in (3.34) and (3.35), we have to find the proper values for scaling and shifting. If $u(n)$ be the modulating signal, and $\hat{u}(n)$ be its estimate, for $n = 1, 2, \ldots, N$, we have to find suitable values for the shift 'p' and scale 'q' such that the expression $\sum_{n=1}^{N} [u(n) - (p + q\hat{u}(n))]^2$ is minimised. The expression for p and q are of the form

$$q = \frac{N\left(\sum_{n=1}^{N} \hat{u}(n)u(n)\right) - \left(\sum_{n=1}^{N} \hat{u}(n)\right)\left(\sum_{n=1}^{N} u(n)\right)}{N\left(\sum_{n=1}^{N} \hat{u}(n)^2\right) - \left(\sum_{n=1}^{N} \hat{u}(n)\right)^2} \tag{3.36}$$

$$p = \frac{\sum_{n=1}^{N} u(n) - q\sum_{n=1}^{N} \hat{u}(n)}{N} \tag{3.37}$$

The error variance of estimation is defined as

$$\frac{1}{N-1}\sum_{n=1}^{N} [u(n) - (p + q\hat{u}(n))]^2.$$

Essentially we estimate the modulating signal via the estimation of $a_1(n)$ and $a_2(n)$. This is done by introducing a suitable set of basis functions for the coefficients, thereby, transforming a non–linear estimation problem into a linear one. The details are given in Appendix A.

As the estimate from the parametric method requires scaling and shifting, these are also utilised for the estimates from the non-parametric methods which are described in the following paragraphs. In this way, we get a fair comparison among their performances of various methods for estimating the modulating signals.

3.2.4 Non–linear Energy Operator

For both continuous and discrete time signals, Kaiser has defined a non–linear energy tracking operator Ψ [52, 53]. For the discrete time case, the energy operator for $x(n)$ is defined as,

$$\Psi[x(n)] \triangleq x^2(n) - x(n-1)x(n+1) \tag{3.38}$$

For the signal

$$x(n) = r(n)\cos(\phi(n)) \tag{3.39}$$

we have

$$\begin{aligned} \Psi[x(n)] &\approx (r(n)\Omega(n))^2 \\ \sqrt{\Psi[x(n)]} &\approx |r(n)\Omega(n)| \end{aligned} \tag{3.40}$$

where $\Omega(n)$ is defined as the instantaneous frequency i.e., $\Omega(n) \triangleq \frac{d\phi(n)}{dn}$. The differentiation operator is approximated with a proper (backward, forward or symmetric) difference operator in the discrete time case.

When one of the variables $r(n)$ and $\Omega(n)$ is constant, we can get the other variable with a scaling of $\sqrt{\Psi[x(n)]}$. So, the energy operator can estimate the modulating signal, or more precisely its scaled version, when either just AM or FM is present [54].

When both AM and FM are present simultaneously, three algorithms are described to estimate $r(n)$ and $\Omega(n)$ separately in [55]. The best among the three algorithms according to performance is called the *discrete energy separation algorithm 1* (DESA-1) [55]. The DESA-1 is defined as follows

$$y(n) = x(n) - x(n-1)$$

$$\arccos\left(1 - \frac{\Psi[y(n)] + \Psi[y(n+1)]}{4\Psi[x(n)]}\right) \approx \Omega(n) \tag{3.41}$$

$$\frac{\Psi[x(n)]}{\sqrt{1 - \left(1 - \frac{\Psi[y(n)]+\Psi[y(n+1)]}{4\Psi[x(n)]}\right)^2}} \approx r(n) \tag{3.42}$$

The frequency estimate works so long as $0 < \Omega(n) < \pi$.

3.2.5 Hilbert Transform Separation Algorithm

The Hilbert transform of any AFM signal $x(n) = r(n)\cos(\phi(n))$ is a signal $\tilde{x}(n)$ with the Fourier transform $\tilde{X}(\omega)$ such that

$$\tilde{X}(\omega) = -j\mathrm{sgn}(\omega)X(\omega) \tag{3.43}$$

where $X(\omega)$ is the Fourier transform of $x(n)$.

Given the analytic signal

$$x(n) + j\tilde{x}(n) = s(n)\exp(j\theta(n)) \tag{3.44}$$

its modulus $s(n)$ and phase $\dot{\theta}(n)$ can serve as (generally approximate) estimates of the amplitude and instantaneous frequency of $x(n)$. Thus the *Hilbert transform separation algorithm* (HSTA) is given by [56]

$$s(n) = \sqrt{x^2(n) + \hat{x}^2(n)} \approx |r(n)| \tag{3.45}$$

$$\theta(n) = \arctan\left(\frac{\hat{x}(n)}{x(n)}\right) \approx \phi(n) \tag{3.46}$$

From the instantaneous phase $\theta(n)$ we can get the estimate of instantaneous frequency by phase unwarping followed up with differentiation or (forward, backward or symmetric) difference in discrete domain.

In practice, Hilbert transform can be implemented by using an FIR or a DFT approximation to the ideal IIR filter [57]. This approximation, however, introduces additional error in the Hilbert transform. The error increases with the decrease in the FIR filter length [56]. For this reason, we have implemented the Hilbert transform with a DFT over the full data length.

3.3 Simulation

3.3.1 Synthesis

For the purpose of simulation we have taken $n_0 = 5000$ so that the signal becomes steady within that time. The origin of the time axis is shifted to 5000 for the purpose of displaying the output. In all the cases, $|a| = 0.99999$ and $\sigma_w = 1.1$ are used for simulation.

For generating the damped sinusoid with $d(m)$ according to (3.10), we have set

$$d(m) = (1.005)^{-m},\ \theta = \frac{\pi}{3};$$

and the real part of the signal as given by (3.11) is plotted in Figure 3.2. In the special case, we set $d(m)$ according to (3.12),

$$d(m) = 0.95 \text{ (a constant), and } \theta = \frac{\pi}{3},$$

and the output sinusoid settles to an AR process with a pole at $0.95\angle\frac{\pi}{3}$ and the real part of output as given by (3.13) is plotted in Figure 3.3.

To generate amplitude modulated signal with $d(n)$ as (3.14), we have chosen

$$d(m) = \exp\left(0.5\sin\left(\frac{\pi}{30}m\right)\right),\quad \theta = \frac{\pi}{3}$$

and the output $real[x(n_0+m)]$ as given by (3.17) is plotted in Figure 3.4. As the $d(m)$ is the exponential of a singletone, the process is named as exponential modulation.

The singletone frequency modulated output is generated with $d(m)$ as (3.20),

$$d(m) = \exp\left(j0.4\sin\left(\frac{\pi}{60}m\right)\right),\ \theta = \frac{\pi}{3}$$

and the output $real[x(n_0+m)]$ as given by (3.23) is drawn in Figure 3.5.

3.3.2 Analysis

Damped Sinusoidal Model

We have generated two damped sinusoids according to (3.9) with :

$$\begin{array}{lll} \text{Case I}: & d(m) = 1.0 - 2.0\times 10^{-5}m, & \theta = \pi/6. \\ \text{Case II}: & d(m) = (1.00001)^{-m}, & \theta = \pi/6. \end{array}$$

for $m = 1, \ldots, 512$. The real part of the two sinusoids are shown in Figure 3.6 and 3.7. In both the cases, we have fitted exponentially damped sinusoids according to Prony's model and the time–dependent AR (TAR) model. For Prony's model, the optimal model orders according to the EDC (2.14) are found to be nine and two in Case I & II respectively. For TAR model, two monomial bases are used in Case I and two Chebyshev bases are employed in Case II. The signal to prediction error ratios are computed for all combinations and shown in Table 3.1 .

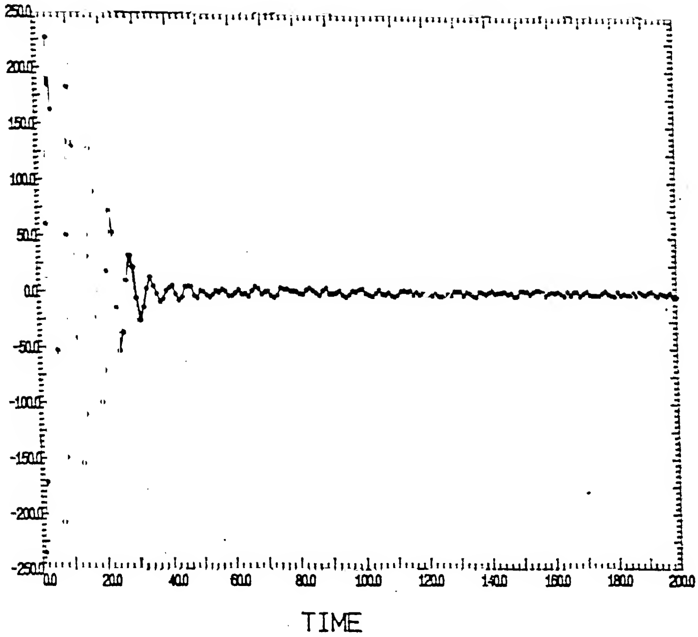


Figure 3.2: The real part of generated damped sinusoid (rapid decay), $Re[x(m)] = \prod_{l=1}^{m} d(l) \cos(\pi m/3)$, $m = 1, \ldots, 200$; and $d(l) = (1.005)^{-l}$.

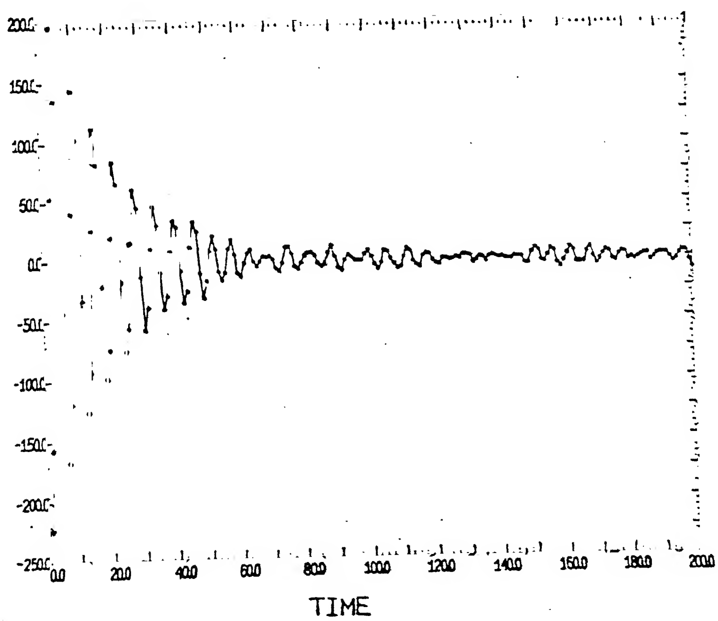


Figure 3.3: The real part of generated damped sinusoid (transient), $Re[x(m)] = \prod_{l=1}^{m} d(l) \cos(\pi m/3)$, $m = 1, \ldots, 200$; and $d(l) = 0.95$.

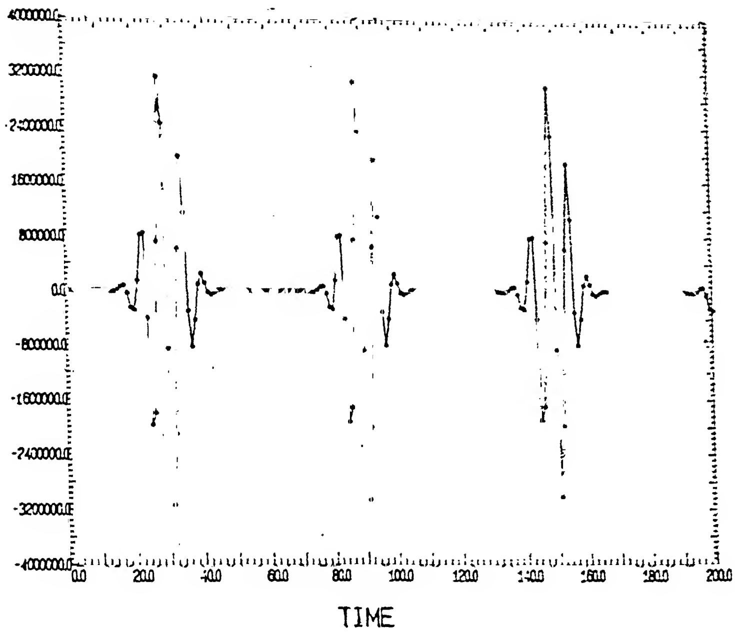


Figure 3.4: The real part of generated amplitude modulated sinusoid (exponential modulation) $Re[x(m)] = \prod_{l=1}^{m} d(l)\cos(\pi m/3)$, $m = 1,\ldots,200$; for $d(l) = \exp(0.5\sin(\frac{\pi}{30}l))$.

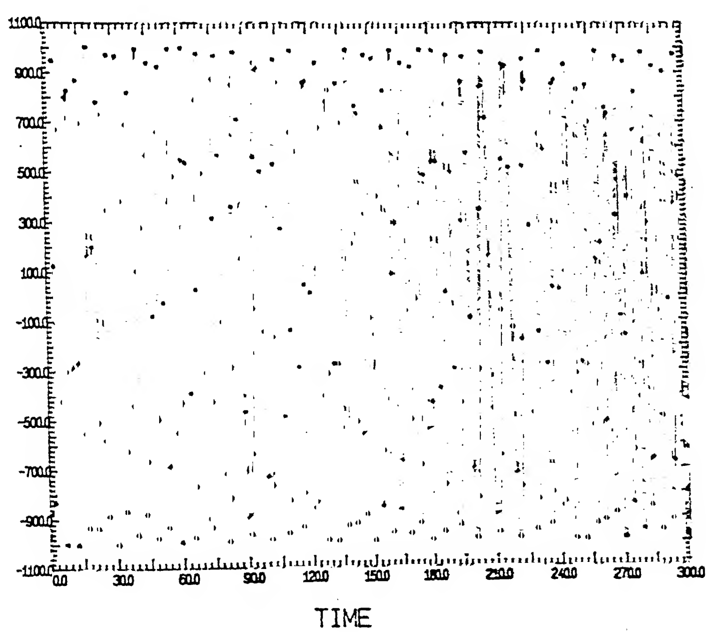


Figure 3.5: The real part of generated frequency modulated sinusoid (single-tone modulation) $Re[x(m)] = \prod_{l=1}^{m} d(l)\cos(\pi m/3)$, $m = 1,\ldots,200$; and $d(l) = \exp(j0.4\sin(\frac{\pi}{60}l))$.

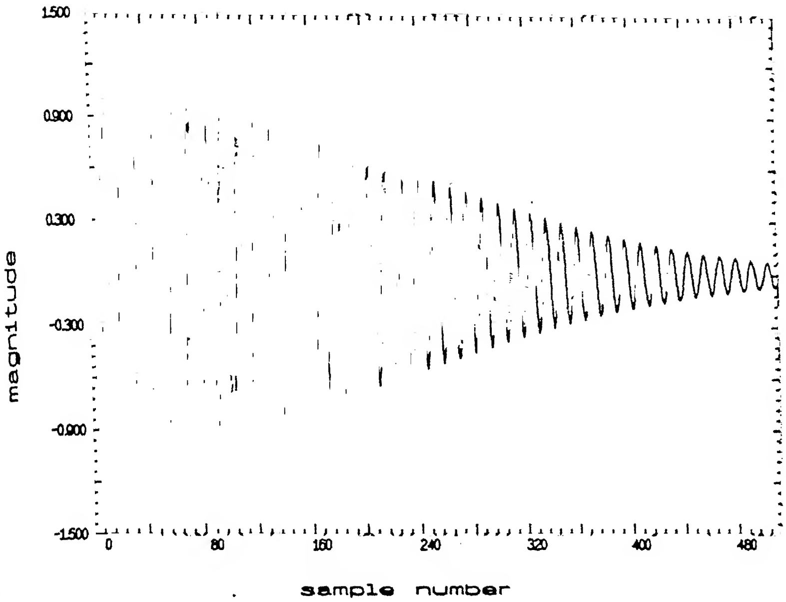


Figure 3.6: The real part of damped sinusoid $Re[x(m)] = \prod_{l=0}^{m} d(l) \cos(\pi m/6)$, $m = 1, \ldots, 512$; and $d(l) = (1.0 - 2.0 \times 10^{-5} l)$.

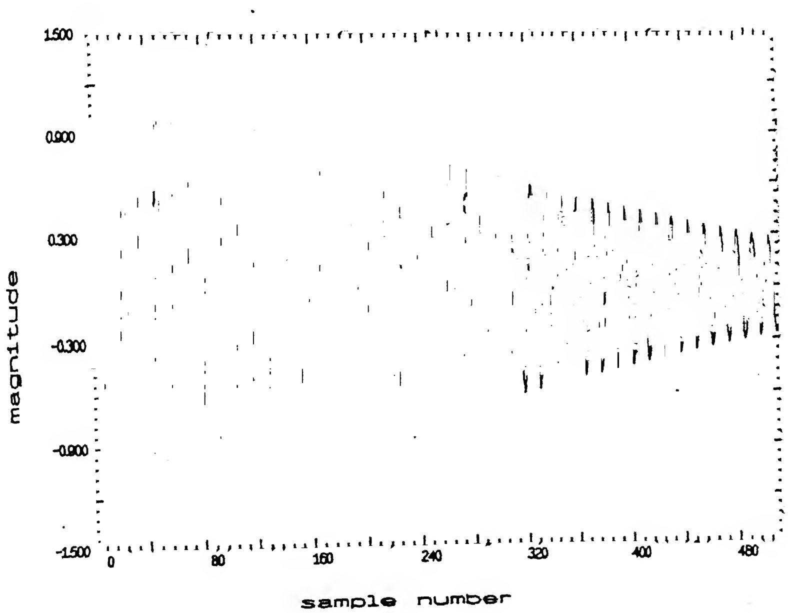


Figure 3.7: The real part of damped sinusoid $Re[x(m)] = \prod_{l=0}^{m} d(l) \cos(\pi m/6)$, $m = 1, \ldots, 512$; for $d(l) = (1.00001)^{-l}$.

Table 3.1: Signal to prediction error ratios in modelling damped sinusoids.

Case	Prony's Model (dB)	TAR (dB)
I	54.5	66
II	58	77

Amplitude Modulated Sinusoidal Model

Consider the discrete time AM signal with singletone modulation, as given by

$$y(n) = (1 + k \cos \omega_a n) \cos (\theta n), \quad n = 1, \ldots, 512;$$

where $k = 0.8$, $\omega_a = \pi/128$ and $\theta = \pi/6$. The modulated signal is shown in Figure 3.8. The modulating signal is estimated as given in Section 3.2.3. While using the Fourier and Chebyshev bases, 33 and 28 basis functions respectively are found to give best results. The estimated modulating signal using the time–dependent AR modêl (TAR) with Chebyshev bases and the error in estimation are shown in Figure 3.9.

The performance is compared to the results obtained using non–parametric approaches like the *non-linear energy operator* and *Hilbert transform*. The DESA-1, which provides the best performence among the three algorithms based on *non-linear energy operators*, is used. For the *Hilbert transform* implementation, a DFT over the full data length is employed for most accurate result [56].The error variance in estimating the amplitude modulating signal by these methods are given in Table 3.2. The loci of the poles are shown in Figure 3.10, which nicely shows the radial movement of the same on the unit circle.

The estimation of AM signal by the proposed method is sensitive to noise. The Figure 3.11 shows the estimates of $a_2(n)$, and the corresponding estimations of the amplitude modulating signals are shown in Figure 3.12, when additive white noise is added to the signal. In order to introduce robustness in the estimation procedure, the system is over–determined using $L = 7$ in (A.12) (Appendix A). In this way, much noise immunity is achieved by considering the innovation at any instant orthogonal to the last L samples or more precisely $[Y(n, 1), \ldots, Y(n, p)]$,

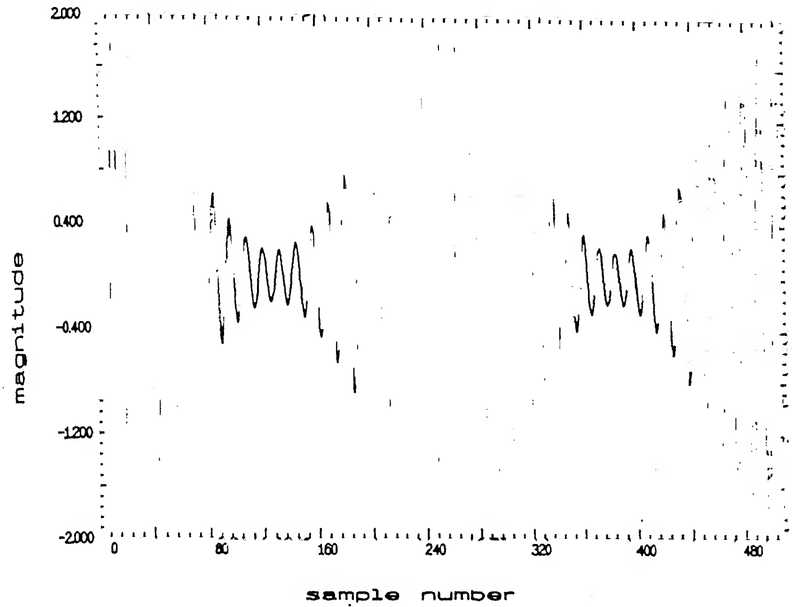


Figure 3.8: The AM signal $y(n) = (1 + k\cos\omega_a n)\cos(\theta n)$, $n = 1, \ldots, 512$; and $k = 0.8$, $\omega_a = \pi/128$, $\theta = \pi/6$.

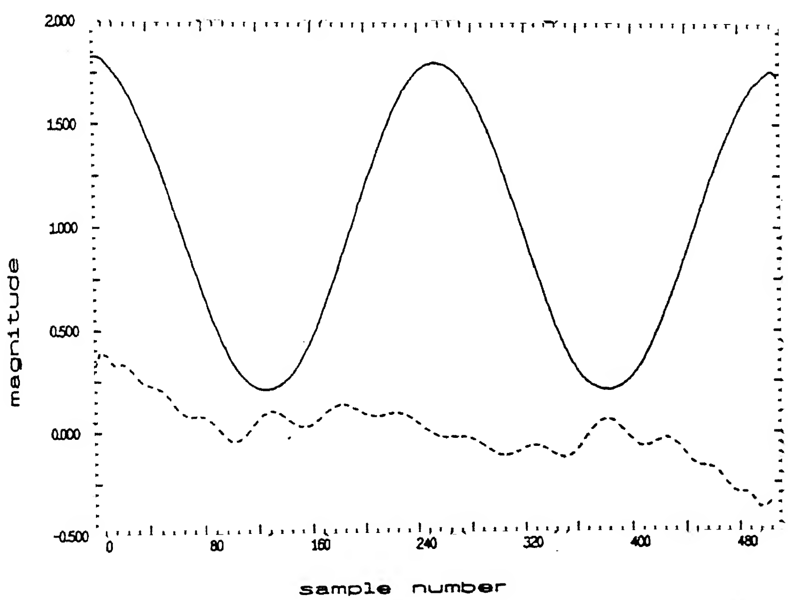


Figure 3.9: The estimated amplitude modulating signal (solid line) and the estimation error (×10) (dotted line).

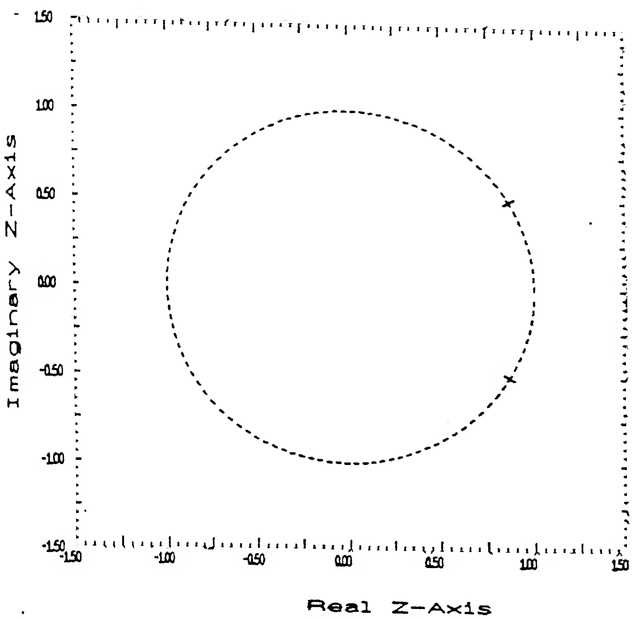


Figure 3.10: The loci of the poles of the AM signal.

Table 3.2: Error variance in estimating the amplitude modulating signal in no noise case.

Non-Parametric Approach		TAR Process implemented with	
Non-linear Energy Operator	Hilbert Transform	Fourier Basis	Chebyshev Basis
1.94×10^{-4}	3.7×10^{-3}	1.39×10^{-4}	4.65×10^{-5}

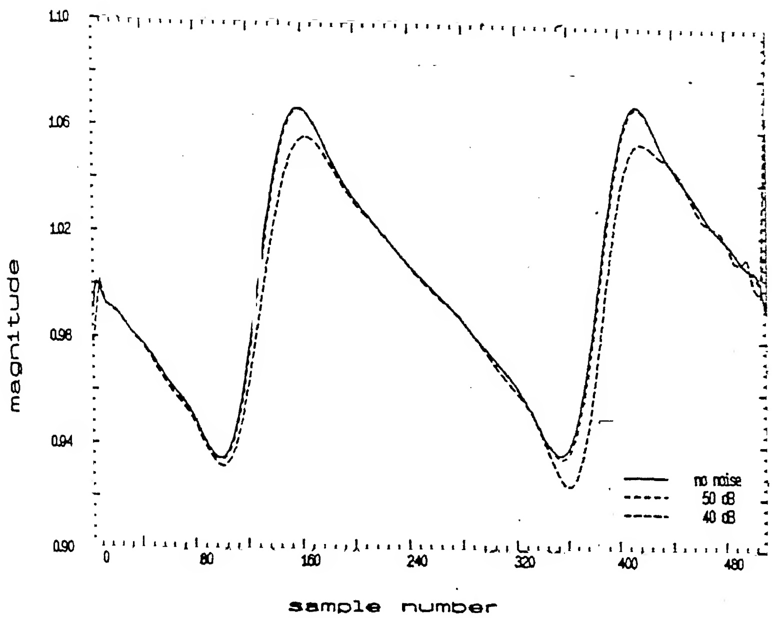


Figure 3.11: The estimates of the time–varying AR coefficient $a_2(n)$ for the AM signal at different SNR levels

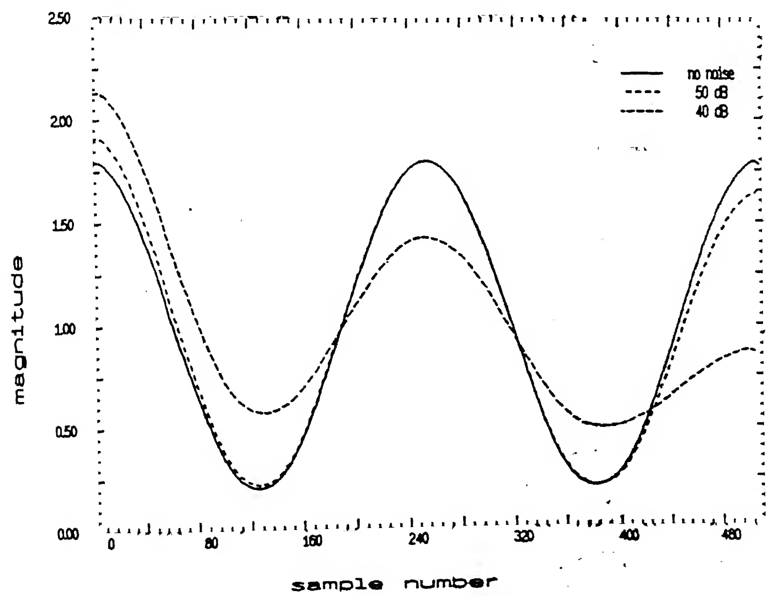


Figure 3.12: The estimates of the modulating signal for the AM signal at different SNR levels

when $L > p$, the AR model order. The additive white Gaussian noise is used for simulation, and 100 realisations are implemented with different noise sequences at each SNR level. The performance of the parametric method in comparison with the performance of the non-parametric methods are shown in Table 3.3. Simulation results for SNR below 40 dB are not shown because in such cases the energy of modulating signal occasionally becomes negative, leading to a meaningless result. This phenomenon is observed while using each of the methods.

Frequency Modulated Sinusoidal Model

Consider the discrete FM signal with singletone modulation,

$$y(n) = cos(\theta n + \omega_m \textstyle\sum_{i=1}^{n} \cos(\omega_f i)), \quad n = 1, \ldots, 512;$$

where $\omega_m = 0.2\theta$, $\omega_f = \pi/128$ and $\theta = \pi/6$. The signal is shown in Figure 3.13. While using the Fourier and Chebyshev bases, 15 basis functions give best results in both the cases. The estimated FM signal in no noise case using Chebyshev basis and the error in estimation are shown in Figure 3.14.

The method for estimating the frequency modulating signal is found to be robust when Gaussian white noise is added to the modulated signal. So, there is no need to make the linear system over-determined in this case. The locus of the poles is shown in Figure 3.15.

The performance of the proposed method in comparison with that of the non-parametric methods are shown in Table 3.4. The results tabulated are obtained by 100 realisations with independent noise sequences at each SNR level. The (*) symbol in Table 3.4 signifies the fact that the energy of modulating signal occasionally becomes negative in some realisations.

Table 3.3: Error variance in estimating the amplitude modulating signal with overdetermined system.

SNR	Non-Parametric Approach		TAR process implemented with	
(dB)	Non-linear Energy Operator ($\mu \pm \sigma$)	Hilbert Transform ($\mu \pm \sigma$)	Fourier Basis ($\mu \pm \sigma$)	Chebyshev Basis ($\mu \pm \sigma$)
∞	1.94×10^{-4}	3.7×10^{-3}	1.22×10^{-4}	7.64×10^{-5}
60	$2.03 \times 10^{-4} \pm 1.51 \times 10^{-6}$	$3.7 \times 10^{-3} \pm 4.14 \times 10^{-6}$	$5.29 \times 10^{-4} \pm 6.01 \times 10^{-4}$	$1.29 \times 10^{-4} \pm 7.29 \times 10^{-5}$
50	$2.76 \times 10^{-4} \pm 6.7 \times 10^{-6}$	$3.71 \times 10^{-3} \pm 1.44 \times 10^{-5}$	$6.84 \times 10^{-4} \pm 9.16 \times 10^{-4}$	$2.62 \times 10^{-4} \pm 7.19 \times 10^{-4}$
40	$3.97 \times 10^{-3} \pm 2.95 \times 10^{-2}$	$3.77 \times 10^{-3} \pm 3.94 \times 10^{-3}$	$1.06 \times 10^{-3} \pm 1.57 \times 10^{-2}$	$2.14 \times 10^{-4} \pm 1.11 \times 10^{-3}$

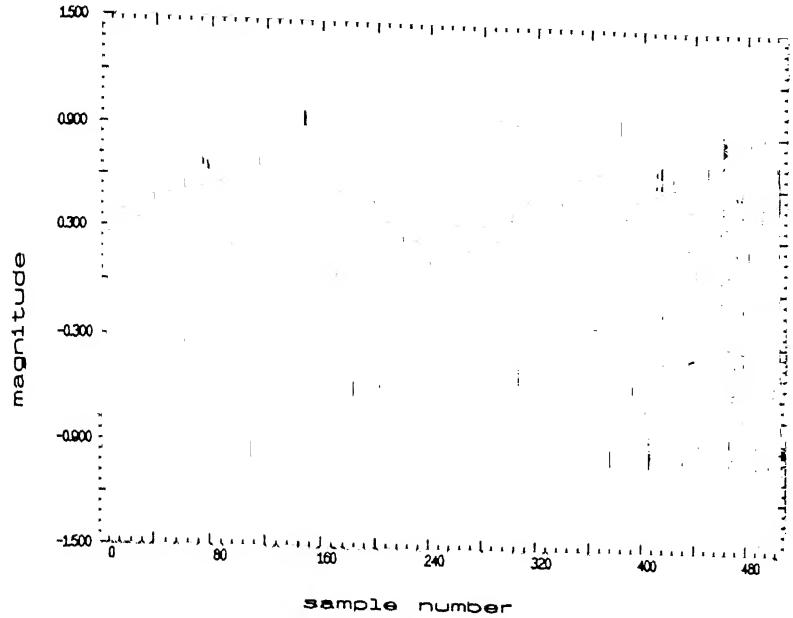


Figure 3.13: The FM signal $y(n) = cos(\theta n + \omega_m \sum_{i=1}^{n} \cos(\omega_f i))$, $n = 1, \ldots, 512$; and $\omega_m = 0.2\theta$, $\omega_f = \pi/128$, $\theta = \pi/6$.

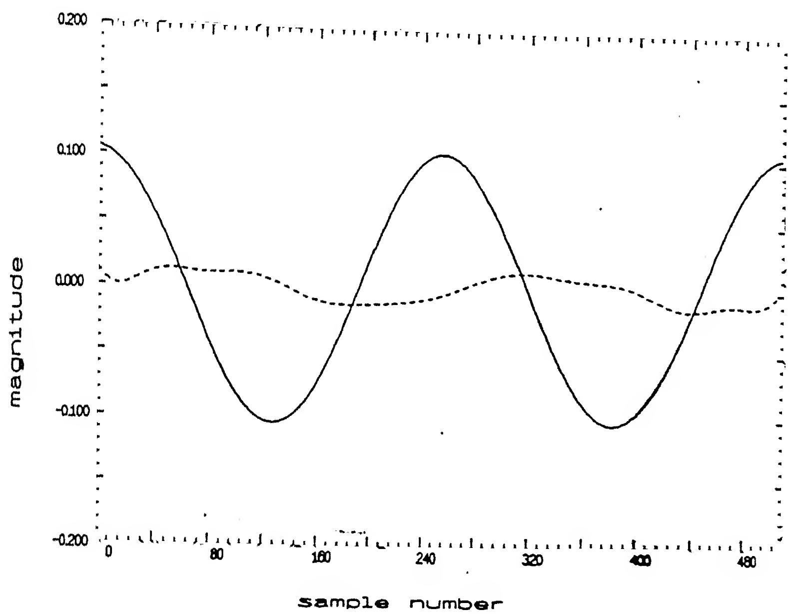


Figure 3.14: The estimated frequency modulating signal (solid line) and the estimation error ($\times 10$) (dotted line).

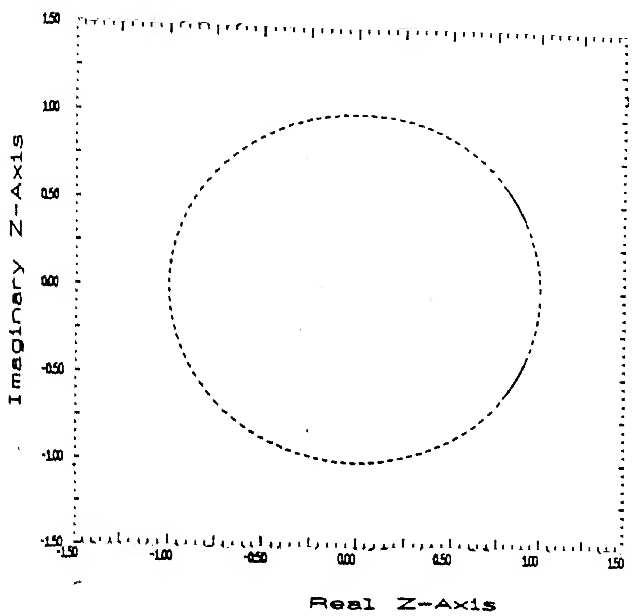


Figure 3.15: The loci of the poles of the FM signal.

Table 3.4: Error variance in estimating the frequency modulating signal.

SNR	Non-Parametric Approach		TAR process implemented with	
(dB)	Non-linear Energy Operator ($\mu \pm \sigma$)	Hilbert Transform ($\mu \pm \sigma$)	Fourier Basis ($\mu \pm \sigma$)	Chebyshev Basis ($\mu \pm \sigma$)
∞	2.65×10^{-6}	1.66×10^{-3}	8.46×10^{-7}	8.79×10^{-7}
60	$5.63 \times 10^{-6} \pm 1.92 \times 10^{-7}$	$1.66 \times 10^{-3} \pm 1.5 \times 10^{-6}$	$8.53 \times 10^{-7} \pm 8.2 \times 10^{-9}$	$8.84 \times 10^{-7} \pm 1.44 \times 10^{-8}$
50	$2.48 \times 10^{-5} \pm 1.48 \times 10^{-6}$	$1.66 \times 10^{-3} \pm 4.75 \times 10^{-6}$	$8.57 \times 10^{-7} \pm 1.58 \times 10^{-8}$	$9.05 \times 10^{-7} \pm 4.62 \times 10^{-8}$
40	$2.07 \times 10^{-4} \pm 1.37 \times 10^{-5}$	$1.68 \times 10^{-3} \pm 1.81 \times 10^{-5}$	$9.4 \times 10^{-7} \pm 9.66 \times 10^{-8}$	$1.18 \times 10^{-6} \pm 2.94 \times 10^{-7}$
30	$1.57 \times 10^{-3} \pm 8.79 \times 10^{-5}$ (*)	$1.93 \times 10^{-3} \pm 4.5 \times 10^{-5}$	$1.59 \times 10^{-6} \pm 7.5 \times 10^{-7}$	$3.55 \times 10^{-6} \pm 2.31 \times 10^{-6}$
20	$4.4 \times 10^{-3} \pm 9.09 \times 10^{-5}$ (*)	$3.37 \times 10^{-3} \pm 8.18 \times 10^{-5}$	$2.07 \times 10^{-5} \pm 1.7 \times 10^{-5}$	$5.04 \times 10^{-5} \pm 4.98 \times 10^{-5}$
15	$5.04 \times 10^{-3} \pm 5.77 \times 10^{-5}$ (*)	$4.42 \times 10^{-3} \pm 5.6 \times 10^{-5}$	$1.13 \times 10^{-4} \pm 6.13 \times 10^{-5}$	$4.31 \times 10^{-4} \pm 8.32 \times 10^{-4}$

Amplitude and Frequency Modulated Sinusoidal Model

Consider the discrete AFM signal,

$$y(n) = (1 + k\cos\omega_a n)\cos\left(\theta n + \omega_m \sum_{i=1}^{n} \cos(\omega_f i)\right), \quad n = 1, \ldots, 512;$$

where, $k = 0.5$, $\omega_m = 0.5\theta$, $\omega_a = \omega_f = \pi/128$ and $\theta = \pi/6$. The signal is shown in Figure 3.16. For the time–dependent AR model while using the Fourier and Chebyshev bases, best results are obtained with 33 and 26 basis functions respectively. The error variances in comparison with that found by non–parametric methods are given in Table 3.5.

Table 3.5: Error variance in estimating the amplitude and frequency modulating signals.

Modulating Signal	Non-Parametric Approach		TAR Process implemented with	
	Non-linear Energy Operator	Hilbert Transform	Fourier Basis	Chebyshev Basis
AM	1.69×10^{-3}	1.41×10^{-3}	3.41×10^{-4}	3.39×10^{-4}
FM	4.68×10^{-4}	2.67×10^{-3}	5.57×10^{-6}	5.75×10^{-6}

The loci of the poles for the AFM process as estimated by the parametric method using the Chebyshev basis are shown in Figure 3.17.

Since estimation of amplitude modulating signal is very sensitive to noise, the system is overdetermined and $L = 7$ is used. For estimation of the modulating signals a total of 32 Fourier bases and 24 Chebyshev bases are employed respectively. The performance of the proposed method in comparison with that of the non–parametric methods is shown in Table 3.6 and 3.7.

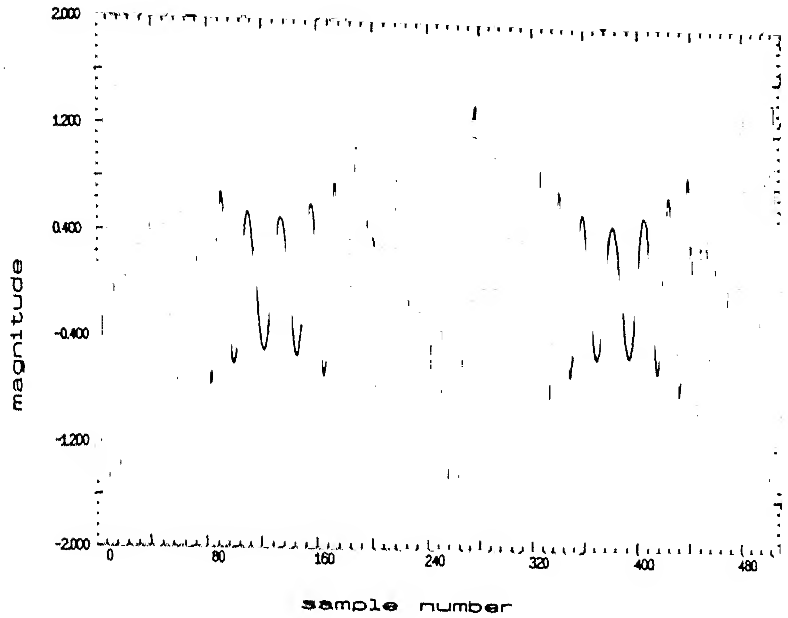


Figure 3.16: The AFM Signal $y(n) = (1 + k\cos\omega_a n)\cos(\theta n + \omega_m \sum_{i=1}^{n} \cos(\omega_f i))$, $n = 1, \ldots, 512$; where, $k = 0.5$, $\omega_m = 0.5\theta$, $\omega_a = \omega_f = \pi/128$, $\theta = \pi/6$.

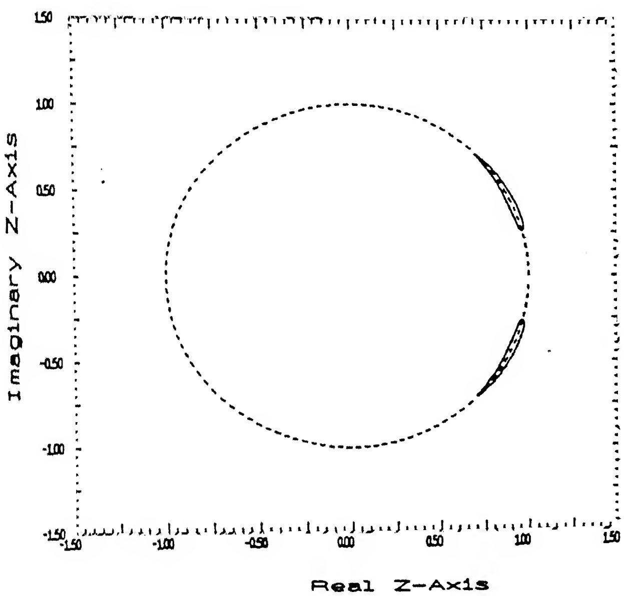


Figure 3.17: The loci of the poles of the AFM signal.

Table 3.6: Error variance in estimating the amplitude modulating signal of the AFM signal with overdetermined system.

SNR	Non-Parametric Approach		TAR process implemented with	
(dB)	Non-linear Energy Operator $(\mu \pm \sigma)$	Hilbert Transform $(\mu \pm \sigma)$	Fourier Basis $(\mu \pm \sigma)$	Chebyshev Basis $(\mu \pm \sigma)$
∞	1.69×10^{-3}	1.11×10^{-3}	3.4×10^{-4}	3.4×10^{-4}
60	$1.69 \times 10^{-3} \pm 1.92 \times 10^{-6}$	$1.13 \times 10^{-3} \pm 2.17 \times 10^{-6}$	$4.92 \times 10^{-4} \pm 1.5 \times 10^{-4}$	$3.41 \times 10^{-4} \pm 4.24 \times 10^{-6}$
50	$1.72 \times 10^{-3} \pm 6.17 \times 10^{-6}$	$1.13 \times 10^{-3} \pm 7.5 \times 10^{-6}$	$5.34 \times 10^{-4} \pm 3.25 \times 10^{-4}$	$3.49 \times 10^{-4} \pm 1.58 \times 10^{-5}$
40	$1.95 \times 10^{-3} \pm 3.59 \times 10^{-5}$	$1.18 \times 10^{-3} \pm 2.24 \times 10^{-5}$	$5.99 \times 10^{-4} \pm 2.90 \times 10^{-4}$	$5.31 \times 10^{-4} \pm 1.64 \times 10^{-4}$

Table 3.7: Error variance in estimating the frequency modulating signal of the AFM signal with overdetermined system.

SNR	Non-Parametric Approach		TAR process implemented with	
(dB)	Non-linear Energy Function ($\mu \pm \sigma$)	Hilbert Transform ($\mu \pm \sigma$)	Fourier Basis ($\mu \pm \sigma$)	Chebyshev Basis ($\mu \pm \sigma$)
∞	4.67×10^{-4}	2.74×10^{-3}	6.46×10^{-6}	5.81×10^{-6}
60	$4.79 \times 10^{-4} \pm 1.46 \times 10^{-6}$	$2.73 \times 10^{-3} \pm 2.72 \times 10^{-6}$	$6.3 \times 10^{-6} \pm 1.02 \times 10^{-6}$	$1.08 \times 10^{-5} \pm 1.11 \times 10^{-5}$
50	$5.64 \times 10^{-4} \pm 1.38 \times 10^{-6}$	$2.74 \times 10^{-3} \pm 7.07 \times 10^{-6}$	$6.35 \times 10^{-6} \pm 1.16 \times 10^{-6}$	$1.14 \times 10^{-5} \pm 4.05 \times 10^{-6}$
40	$2.93 \times 10^{-3} \pm 3.31 \times 10^{-3}$	$2.84 \times 10^{-3} \pm 2.84 \times 10^{-5}$	$7.0 \times 10^{-6} \pm 1.27 \times 10^{-6}$	$2.2 \times 10^{-5} \pm 1.85 \times 10^{-5}$

3.4 Discussion

Parametric modelling of non–stationary signals by the time–dependent ARMA (TARMA) model is investigated here. It is shown that the approach is generalised, and modelling by complex exponentials, AFM, etc. signals are special cases under the general framework. This fact is confirmed by both synthesis and analysis of different type of signals. In the general framework of TARMA model, the damped sinusoid, AM, FM and AFM models can be classified by the positions/loci of the poles of the process. In case of the damped sinusoidal model, the poles are located within the unit circle. When the damping factor is constant, the poles are located at fixed points, whereas for time–varying damping the loci of the poles are given by the radial lines. For the AM signal model, the loci of the poles are the radial lines intersecting the unit circle as shown in Figure 3.10 . For the FM signal model, the poles have only the circumferential movements on the unit circle as shown in Figure 3.15. For the AFM signal model, both the radial and circumferential movements of the poles are observed and the loci of the poles intersect the unit circle as shown in Figure 3.17.

Transient signals like nuclear magnetic resonance (NMR) [21] and electromagnetic pulse (EMP) [58, 45] responses are traditionally represented by the complex exponential models. Example 1 of section 3.3.2 (analysis) shows that an arbitrarily damped sinusoid can be represented more suitably by the time–dependent AR (TAR) model, with a damping factor as a function of time. The TAR model is found to be more flexible, and it gives a new way of accurate modelling of transient signals.

As shown in the examples of section 3.3.2 (analysis), the parametric method of the TAR model clearly outperforms the non–parametric methods of estimating the modulating signals in the no-noise case. In case of a signal corrupted with additive white noise the parametric method excels in performance when suitable basis functions are used. The estimation of frequency modulating signals is found to be more robust than the estimation of amplitude modulating signals, when corrupted with additive white noise. This is because in case of amplitude modulation any error in estimating $a_2(n)$ at any instant 'n' gets accumulated as shown in (3.32) and (3.34).

Conventionally, the speech like signals are modelled by the AM/FM signals for

the purpose of analysis/ synthesis [46, 47, 48]. The presence of the AM and FM signals in speech production was reported earlier [28]. Recently the experiments on speech resonance also indicate the presence of the AM and FM signals significantly [55, 56]. The FM signal is frequently encountered in sonar applications [59]. In underwater environment, the rotating machinery (viz. submarine propeller etc.) produces the AFM signal [54]. These phenomena motivate us to develope the parametric methods for estimating the AM and FM signal with greater accuracy. Another point to be mentioned here is that unlike the non-parametric methods, the parametric approach can be extended to the case of multiple AM or FM sinusoids. Preliminary results indicate the potential of the parametric method and further work is needed to use these techniques in real life situations.

It should be emphasised that in this thesis we have tried to find the relations among various existing non–stationary models and exploit the relations for better modelling and estimation. For the purpose of estimation of model parameters, however, we have relied on the techniques which are already available. The shortcomings of the estimation procedures are discussed in section 5.2 in detail.

Chapter 4

ECG Signal Synthesis by Exponential AM Sinusoids

4.1 Introduction

The Electrocardiograph (ECG) is a non–stationary signal [2]. The burst–like QRS feature attributes localised high frequency components in the ECG signal, making it distinctly non–stationary in nature [60]. Though this feature of QRS wave has helped detection of the wave by filtering/linear prediction [29], it makes the modelling of the signal very difficult.

Most of the work in modelling ECG are non–parametric in nature [61, 62]. It is well known that parametric modelling is superior to non–parametric modelling for improved accuracy and flexibility in analysis and synthesis. The improvement results because more information is provided about the signal [1].

Previous attempt to represent parametrically a segment of the ECG either by the impulse response of a pole–zero model or by Prony's model was not very successful because the models had not considered explicitly the non–stationarity of the modelled signal or its pseudo-periodicity [63, 64]. This led to prohibitively large model orders, or more seriously, the model could represent only the local characteristics of the ECG signal. It has been, therefore, the main motivation of the present thesis to develop a parametric model for the ECG signal, which will represent its global nature.

The time–dependent AR/ARMA model is the representative of the general class of non–stationary signals [2]. As such the model can be used for the ECG signal as well. However, the ECG has some distinctive features viz. i) its pseudo–periodicity;

ii) different features of the constituent signals (P, QRS and T) representing actions of various parts of the heart [30]; etc. It will be insightful to know how the general time–dependent AR/ARMA model is restricted by the special features of the ECG–type signals.

It is shown in this thesis that the amplitude modulated (AM) sinusoidal signal model is a special case of the time–dependent AR/ARMA model. The AM signal model can have the periodicity property, and can exhibit burst–like feature very well, when the modulating signal is an exponential function. Hence, it is proposed that one or more AM sinusoidal signal(s) can be employed to model separately each constituent signal of the ECG. The suitability of the developed model is then investigated for the ECG signal using analysis–by–synthesis technique.

4.2 Study of the Model

4.2.1 Model Development

In Section 3.2.2, it is shown that from a single order AR process (3.6) :

$$x(n) = \begin{cases} ax(n-1) + w(n) & \text{if } n < n_0 \\ ad(n)x(n-1) + w(n) & \text{if } n \geq n_0 \end{cases} \tag{4.1}$$

with $a \approx 1.0 \exp(j\theta)$, $d(n) = \exp(b \sin \alpha(n - n_0))$ $n \geq n_0$ and the noise variance σ_w^2 approaching zero, one can get the sequence

$$\begin{aligned} y(m) &= Real\,[x(n_0 + m)] \\ &\simeq A \exp\left(\frac{b}{\alpha}(1 - \cos \alpha m)\right) \cos(\theta m + \phi) \end{aligned} \tag{4.2}$$

where $x(n_0) = A \exp(j\phi)$. The sequence $y(m)$ is an amplitude modulated (AM) sinusoid with carrier frequency θ and modulating factor given by $\exp\left(-\frac{b}{\alpha} \cos \alpha m\right)$. We term the signal $y(m)$ as outcome of exponential amplitude modulation (EAM). The presence of exponential form helps in generating the burst–type periodic signals like the ECG.

Analysis–by–Synthesis Rules

The parameters of the signal model can be determined through the estimation of process parameters, as suggested in Reference [2]. A suitable set of basis functions

can be used to express the time-varying process parameters. We will, however, adopt here an analysis–by–synthesis approach for finding the parameters of the ECG signal directly. The approach will provide some insight regarding the composition of the signal.

In the analysis–by–synthesis approach, first note that the carrier signal frequency θ must be an integer multiple of the modulating frequency α to have similar waveshape in every period of the signal $y(n)$. Though the ECG signal is only pseudo–periodic, we have considered the ideal case where the modelled signal becomes a periodic extension of one period of the ECG signal. It is done to make the synthesis procedure simple. Otherwise, making θ and α time–varying while keeping their harmonic relationship intact, the actual ECG data can also be modelled.

Under the condition of strict periodicity for $y(m)$, (4.1) can be rewritten as ,

$$y(m) = A \exp\left(\frac{b}{\alpha}(1 - \cos \alpha m)\right) \cos(k\alpha m + \phi) \tag{4.3}$$

where k is an integer.

To generate a complicated signal pattern more than one AM sinusoids are needed. The most prominent feature of the signal is first fitted by one AM wave, and once fitted, this part of the original signal is removed. We continue this process on the residual signal till all the features of the signal under investigation is exhausted. For the purpose of fitting an EAM sinusoid, the following steps are employed :

1. We start with some initial guess of A, b and k. At the starting $\phi = 0$, and $\alpha = \frac{2\pi}{T}$, where T is the period of the signal under investigation.

2. The value of the parameters b and k are to be altered to get proper shape of the wave. The phase ϕ is also changed to decide the sign of the peak and skewness of the hump. So, the parameters b, k and ϕ are changed iteratively till the shape of the synthesised signal comes close to the target feature of the ECG signal e.g. the QRS complex.

3. Then, the signal is scaled by a factor c and the origin is shifted to m_0, so that the target feature of the original signal coincides with the synthesised one. If the fitting is not satisfactory at this stage, we have to go back to step 2 and retry. The synthesised signal $z(m)$ is expressed as,

$$z(m) = c\, y(m + m_0) \tag{4.4}$$

4. Once the synthesised signal is ready, it is subtracted from the signal under investigation. In this way, we get the residual signal. If the residual signal still has some prominent feature, we have to go to step 1 to start with a new EAM sinusoid. When all features of the signal are exhausted, the process is stopped.

4.2.2 ECG Signal Fitting

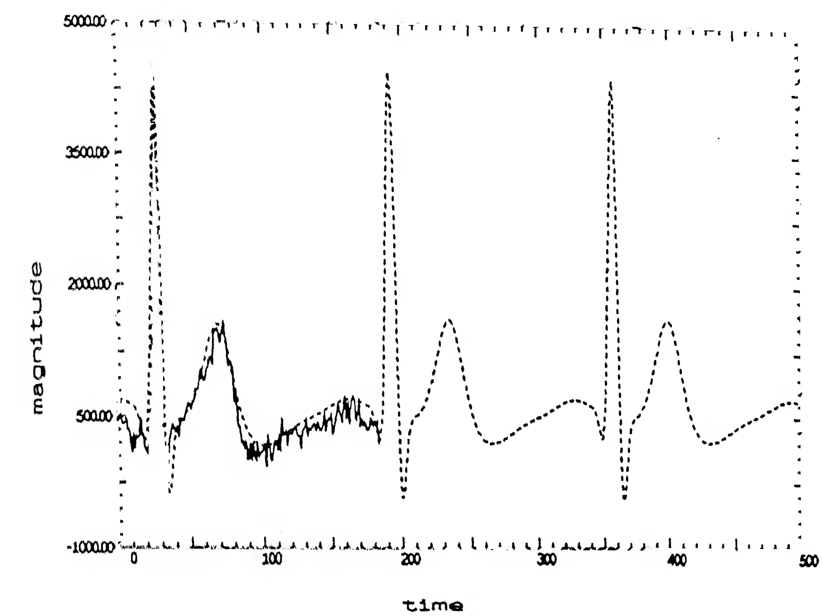


Figure 4.1: The original ECG signal, Lead II configuration, 180 samples sampled at 250 Hz (solid), and synthesised ECG signal (dash).

The Lead II ECG signal of a normal male patient collected for one period and sampled at 250 Hz is shown in Figure 4.1. The T and QRS waves are synthesised by two separate AM sinusoids. For P wave, two more AM sinusoids are needed, and the corresponding waves are named as P1 and P2. A d.c. signal is also needed to match the d.c. bias of the original ECG signal. The successive waves are matched in the following order : T, QRS, P1, and then P2. These waves are shown in Figures 4.2 – 4.3. While fitting the T wave, the d.c. bias of the signal is adjusted.

To synthesise the ECG signal components, we start with $|a| = 0.9999$, $\sigma_w = 10^{-28}$ and $n_0 = 15000$, as defined in (4.1) . The d.c. bias is found to be 500. The values of the other parameters for the constituent waves T, QRS, P1 and P2 are given in

4. Once the synthesised signal is ready, it is subtracted from the signal under investigation. In this way, we get the residual signal. If the residual signal still has some prominent feature, we have to go to step 1 to start with a new EAM sinusoid. When all features of the signal are exhausted, the process is stopped.

4.2.2 ECG Signal Fitting

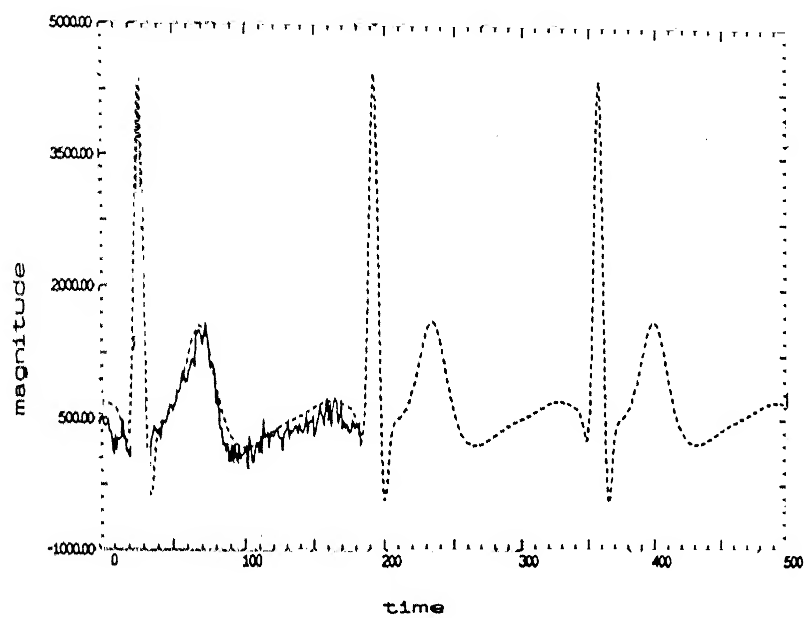


Figure 4.1: The original ECG signal, Lead II configuration, 180 samples sampled at 250 Hz (solid), and synthesised ECG signal (dash).

The Lead II ECG signal of a normal male patient collected for one period and sampled at 250 Hz is shown in Figure 4.1. The T and QRS waves are synthesised by two separate AM sinusoids. For P wave, two more AM sinusoids are needed, and the corresponding waves are named as P1 and P2. A d.c. signal is also needed to match the d.c. bias of the original ECG signal. The successive waves are matched in the following order : T, QRS, P1, and then P2. These waves are shown in Figures 4.2 – 4.3. While fitting the T wave, the d.c. bias of the signal is adjusted.

To synthesise the ECG signal components, we start with $|a| = 0.9999$, $\sigma_w = 10^{-28}$ and $n_0 = 15000$, as defined in (4.1) . The d.c. bias is found to be 500. The values of the other parameters for the constituent waves T, QRS, P1 and P2 are given in

the following table.

Table 4.1: Parameter values for the ECG signal synthesis

Parameters	T wave	QRS wave	P1 wave	P2 wave
α	0.03808^c	0.03808^c	0.03808^c	0.03808^c
b	0.19	1.1	0.1	0.1
k	2	7	1	1
ϕ	$0.25\pi^c$	$1.15\pi^c$	$1.6\pi^c$	π^c
m_0	5	54	155	85
c	0.78557	4.778×10^{-21}	40.0	15.0
A	7.258×10^{-2}	6.917×10^{-2}	2.505×10^{-2}	2.505×10^{-2}

The synthesised ECG signal is shown in Fig 4.1, which closely resembles the original signal, over the period of consideration. As the synthesis technique has produced a periodic extension of one period of the ECG signal, we compute SNR and correlation coefficient over the period concerned. The SNR is calculated to be 9.06dB, when SNR is defined as,

$$\mathrm{SNR} = 10.0 \log \frac{\sum_{i=1}^{n} (x_{org}(i))^2}{\sum_{i=1}^{n} (x_{org}(i) - x_{syn}(i))^2}$$

where $x_{org}(i)$ and $x_{syn}(i)$ are the original and synthesised ECG signal at i-th instant, and n is the period of the ECG signal (= 165 samples). The correlation coefficient (ρ) between the synthesised and original ECG is calculated to be 0.933. The correlation coefficient is defined by

$$\rho = \frac{\sum_{i=1}^{n}(x_{org}(n) - m_{org})(x_{syn}(n) - m_{syn})}{\sqrt{\sum_{i=1}^{n}(x_{org}(n) - m_{org})^2 \sum_{i=1}^{n}(x_{syn}(n) - m_{syn})^2}}$$

where n, $x_{org}(i)$ and $x_{syn}(i)$ are as defined earlier, and m_{org} and m_{syn} are the mean of original and synthetic ECG signal respectively. The generating process can be represented as shown in Fig 4.4. Equivalently, Fig 4.5 provides the complete process representation of the signal. Note that the functional block, Real[$\cdot$] does not appear in the latter diagram.

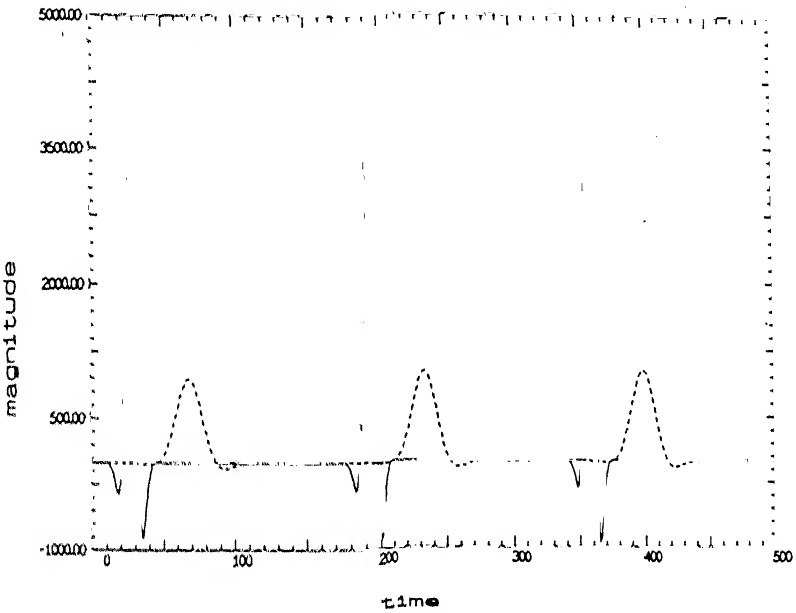


Figure 4.2: The synthesised QRS wave (solid), and synthesised T wave (dash).

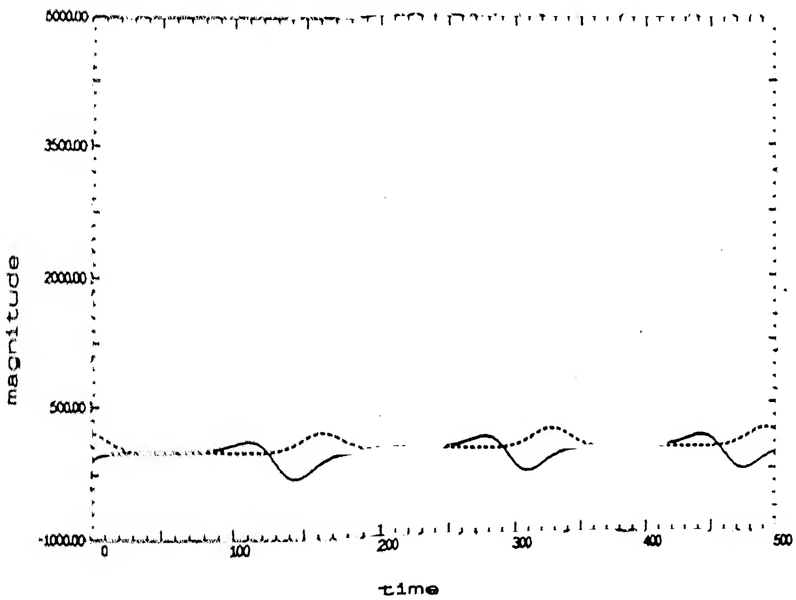


Figure 4.3: The synthesised components of P wave, P1 (solid) and P2 (dash).

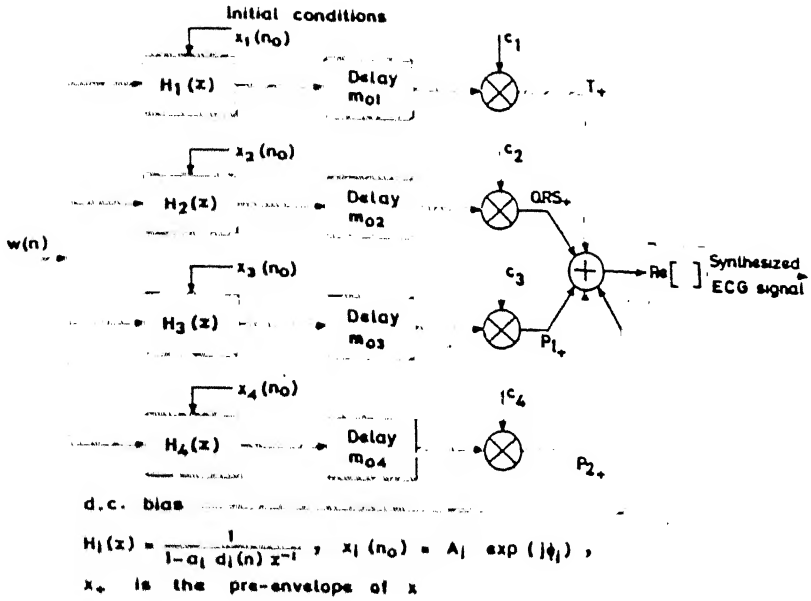


Figure 4.4: The block diagram of the ECG signal synthesiser by first order complex AR models.

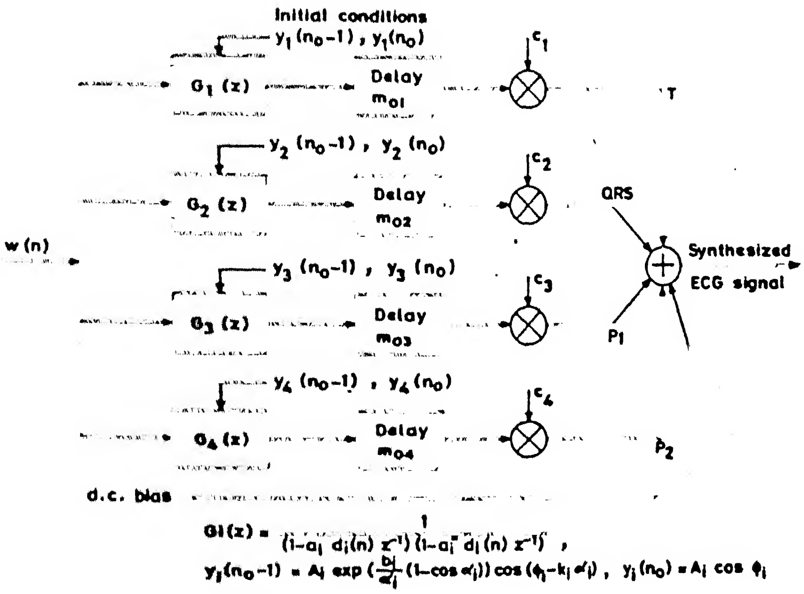


Figure 4.5: The block diagram of the ECG signal synthesiser by second order real AR models.

4.3 Remarks

The method presented here provides a novel way to parametrically model the ECG signal, and also its constituent waves. It is known that the P and QRS waves are caused by the electric potentials generated as the depolarisation prior to the atria and ventricles respectively. Whereas, the T wave in the ECG is caused by the potential generated as the repolarisation of the ventricles. Therefore, it is meaningful to model the constituent waves separately as it is done in the present approach, because they correspond to the separate actions of the various parts of the heart [30].

It is apparent that the ECG generating model as shown in Figure 4.4 or 4.5 gives an equivalent process representation of various functions of the heart. As almost all the serious abnormalities of the heart can be detected by analysing the ECG wave from different leads, this model can be expected to provide insight in the functioning of the heart, and may be helpful for diagnosis [30]. It is clear however that considerable amount of further research will be needed in this direction before the model can be employed for clinical diagnosis.

Until now, the data compressions for the ECG are implemented mostly by non–parametric approaches because the parametric methods, in most cases, did not preserve the waveform [65]. The method presented here being a synthesis technique, guarantees the preservation of waveform. Moreover, the use of a small number of parameters, (only 25 parameters required) promises the application of the model for superior data compression, but it can only be asserted after further research.

For the analysis of the signal, direct evaluation of model parameters leads to a set of highly nonlinear equations. As an alternative, one may estimate the equivalent process parameters by the method as shown in Appendix A. The equivalence between the two approaches as shown in this thesis is an important concept. The concept can be conveniently utilised for parameter estimation of the signal model.

The analysis–by–synthesis approach as demonstrated here is simple to implement. The approach clearly provides good insight about modelling of the ECG signal [30].

Chapter 5

Conclusion & Scope for Future Work

5.1 Conclusion

In the study of TARTG model, we have found that both the time-varying AR coefficients and the time-varying gain are important in their own ways. The first set represents the spectral waveshape variation over time and the latter keeps the information of time variation of the spectral gain. Obviously, none can represent the information of the other. The estimation of gain allows us to obtain a faithful estimation of the evolutionary spectrum, and in this way, the interpretation of the evolutionary spectrum as the energy density in the time-frequency plane becomes apparent.

In parametric modelling, the parsimony of the model is an important issue. If a model requires a large number of parameters, then it is clear that the model is not properly fitting into the signal [66]. In this case, search should be continued for more suitable model. In modelling the speech phonemes, we have found that the QSAR and TAR models require a large number of parameters compared to the ARTG model, while the signal-to-prediction error ratio varies only slightly. So, the ARTG is a more suitable model for speech phonemes, which can be directly used in speech synthesis. In case of speech coding, often the QSAR model is used, and thereafter, the redundancy of parameters are removed by different vector quantisation techniques [35]. If in continuous speech, the phoneme boundaries or more precisely, segmentation of frames with different spectral waveshapes can be done, the ARTG model can be used for speech coding. The alternative model can save

a lot of computation for searching of the code-book. In the same way, the ARTG with proper segmentation will be a better choice for linear prediction (LP) based speech recognition algorithm.

Even in case of the TAR model (with constant gain) when used for coding, the estimation of gain should be used for normalisation. This can be illustrated by an example as follows. Let us consider two different stationary AR processes of different signal strengths placed one after another. When a TAR model is fitted for the combined sequence using a finite number of basis functions, the obvious result of estimation of parameters by minimising the unweighted error function will be a close fit in the portion where the signal strength is high, and will not follow the portion of the signal where the signal strength is low. As in a TAR model any frame under consideration can have vowels and consonants, the effect will be similar if the signal is not normalised using the estimated time-varying gain.

The investigation on the movement of poles of the time-dependent RTF model has revealed that the damped sinusoidal model, AM sinusoidal model, FM sinusoidal model or AM/FM sinusoidal model are only special cases of the generalised model. Each of these models can be generated by putting proper restrictions on the poles of the time-dependent RTF model. This knowledge gives us a unified procedure for estimation of parameters. This leads to more flexibility/accuracy of modelling of damped sinusoids and better accuracy in estimating the modulating signals in case of modulation models.

The last part of the thesis is devoted to modelling of the ECG signal. The synthesis technique presented has given insight about the ECG modelling. The developed model is promising and it may be useful in explaining various functions/malfunctions of the heart.

5.2 Scope for Future Work

The study on parametric modelling of non–stationary signals as shown in this thesis has shown great potential for various applications in future. Before harnessing the advantage, however, a few hurdles have to be crossed. In the present work of unification of various linear non-stationary models and classification of non–stationary signals, the 'standard' method of estimation through basis functions is used for each AR coefficient of the TAR model. The method, however, has several drawbacks and some of those are demonstrated in section 3.3.2 (Alalysis). Therefore, one has to

find suitable solutions to a number of issues before the modelling technique can be applied in real–life solutions. The issues are :

- Choice of the basis function
- Number of the basis functions to be used
- Order of the model

Although these issues are addrssed separately, these are highly dependent on one another. The choice of basis function refers to what should be the natural choice of basis. For example, when the time–variation of the parameter is linear, the sinusoidal basis is quite inappropriate and its choice may lead to a very large basis order. So, before employing any set of basis function, one must have an idea about the time–variation pattern of the parameter under consideration. The issue of optimal basis order selection arises next when a set of basis functions is selected. The order of the model and the number of basis functions should be employed are also inter-dependent. As the error in fitting the model (i.e. prediction error in case of the AR model) can be minimised either by increasing the number of bases or by increasing the model–order, the optimal choice of these quantities are very difficult. The question may be complicated for a multi–component signal where only one component is prominent at a time, e.g. the ECG. Apperently, we can have a large model–order with a small number of basis functions or a small model–order with a large number of basis functions. So, the conventional root-mean-square error judgement criterion can not be expected to give a prudent answer to this question. For this purpose, some new approach and new theory should be developed for non–stationary signals. This should be done while utilising the information about the production/perception mechanism of the signal in concern. Once these questions are reasonably answered, attention can be diverted to make the algorithms more robust from noise. This remains to be the motivation for future research and the suggestion for the future researchers in this field.

Appendix A

Estimation of Time–Dependent AR Coefficients

The time-dependent autoregressive (TAR) model is represented by the equation,

$$x(n) + a_1(n)x(n-1) + \cdots + a_p(n)x(n-p) = w(n) \tag{A.1}$$

for $n = 1, 2, \ldots, N$; where p is the autoregressive model order, $w(n)$ is the white noise sequence and, $a_i(n)$ is the i-th AR coefficient at an instant 'n'.

The coefficient $a_i(n)$ is assumed to be 'smooth' in the sense that if the first derivative of the coefficient may be arbitrarily great, the higher order derivatives necessarily vanish. So, the coefficients $a_i(n)$ can be approximated by a set of basis functions. The time–dependent coefficient $a_i(n)$ is expressed as,

$$a_i(n) = \sum_{j=0}^{m} a_{ij} f_j(n) \tag{A.2}$$

where $f_j(n)$ is the value of j th basis function at an instant 'n'. Let us consider the quantity $a_i(n)x(n-i)$ which can be written as,

$$\begin{aligned} a_i(n)x(n-i) &= \left(\sum_{j=0}^{m} a_{ij} f_j(n)\right) x(n-i) \qquad \text{(A.3)} \\ &= \sum_{j=0}^{m} a_{ij}\left(f_j(n)x(n-i)\right) \end{aligned}$$

Let us define the vector $X(n,i) = \left[f_0(n)x(n-i), \ \ldots, \ f_m(n)x(n-i) \right]^T$, where symbol '$T$' stands for transpose of a vector. With this definition $(A.3)$ becomes

$$a_i(n)x(n-i) = \left[a_{i0} \ \cdots \ a_{im} \right] X(n,i) \tag{A.4}$$

Using $(A.4)$, the TAR process $(A.1)$ can be written as

$$x(n) + \begin{bmatrix} X^T(n,1) & \cdots & X^T(n,p) \end{bmatrix} \theta = w(n) \tag{A.5}$$

where

$$\theta = \begin{bmatrix} a_{10} & \cdots & a_{1m} & a_{20} & \cdots & a_{2m} & \cdots & a_{p0} & \cdots & a_{pm} \end{bmatrix}^T \tag{A.6}$$

The interpretation of $(A.5)$ is immediate. The parameter vector θ is the vector of regression of the non–stationary process $x(n)$ on the past samples of the vector $X(n,i)$ for $i = 1,\ldots,p$. Here lies the advantage of basis parameterisation : a linear time varient or non–stationary problem becomes a linear time-invariant or stationary problem by replacing a scalar process by a vector one.

The innovation process $w(n)$ of the TAR can also be viewed as a prediction error

$$w(n) = x(n) - \hat{x}(n) \tag{A.7}$$

where $\hat{x}(n) = E(x(n)|x(n-1),\cdots,x(n-p))$ is the prediction of $x(n)$ given the past values $x(n-1),\ldots,x(n-p)$, i.e.

$$\hat{x}(n) = -\begin{bmatrix} X^T(n,1) & \cdots & X^T(n,p) \end{bmatrix} \theta \, . \tag{A.8}$$

It is therefore meaningful to minimise the variance of this error $w(n)$ and hence we get [2],

$$E\left[(\text{grad}_\theta w(n)).w(n)\right] = 0 \tag{A.9}$$

The gradient can be easily computed as

$$\text{grad}_\theta w(n) = \begin{bmatrix} X^T(n,1) & \cdots & X^T(n,p) \end{bmatrix}^T \tag{A.10}$$

We therefore obtain the optimum vector θ as the solution of the Yule–Walker type equation,

$$E\left(\begin{bmatrix} X(n,1) \\ \vdots \\ X(n,p) \end{bmatrix} \begin{bmatrix} X^T(n,1) & \cdots & X^T(n,p) \end{bmatrix} \right) \theta = -E\left(\begin{bmatrix} X(n,1) \\ \vdots \\ X(n,p) \end{bmatrix} x(n) \right) \tag{A.11}$$

Remark 1: If the coefficients are constants i.e. $m = 0$ and $f_0(n) = 1$, then $(A.11)$ becomes the Yule–Walker equation of AR(p) process.

Remark 2: If the process $x(n)$ is purely auto–regressive but, $w(n)$ has a gain $g(n)$, it is no longer significant to minimise the prediction error. However, $(A.11)$ remains valid in the sence it expresses nothing but the orthogonality of the innovation $w(n)$ and this orthogonality still makes sense, even if $g(n)$ is not constant [2].

As $(A.11)$ is obtained through minimisation of the variance $w(n)$, which is characterised as a stationary and white noise process, it is possible to use in $(A.11)$ the ergodic estimator for expectation, even if the original process $x(n)$ is non–stationary. This allows us to replace the expectation by a summation over a time interval $\tau = [p+1, N]$ [2].

When the signal is embedded in additive white noise, an overdetermination in the matrix equation $(A.11)$ can give better estimate [32]. In this case, $(A.11)$ takes the form

$$E\left(\begin{bmatrix} X(n,1) \\ \vdots \\ X(n,p) \\ \vdots \\ X(n,L) \end{bmatrix}\begin{bmatrix} X^T(n,1) & \cdots & X^T(n,p) \end{bmatrix}\right)\theta = -E\left(\begin{bmatrix} X(n,1) \\ \vdots \\ X(n,p) \\ \vdots \\ X(n,L) \end{bmatrix} x(n)\right) \tag{A.12}$$

where $L > p$. The extra set of equations results from the fact that the innovation $w(n)$ at an instant 'n' is also orthogonal to the past samples $X(n, p+1), \ldots, X(n, L)$. In this case, the time interval τ over which the summation replaces the expectation becomes $[L+1, N]$.

For the purpose of representing the time–variation of the AR coefficients, the Fourier and Chebyshev basis functions are employed in this thesis. The k-th Fourier basis is defined as

$$f_k(n) = \begin{cases} \cos\left(\frac{\pi(n-1)k}{N-1}\right) & \text{for even value of k,} \\ \sin\left(\frac{\pi(n-1)k}{N-1}\right) & \text{for odd value of k,} \end{cases}$$

where $1 \le n \le N$. The k-th Chebyshev basis is defined as

$$f_k(n) = \cos\left(k \arccos\left(l-1\right)\right),$$

where $l = \frac{2(n-1)}{N-1}$ and $1 \le n \le N$.

Appendix B

Fourier Cosine Series with Half–Range Expansion

Let a continuous time signal $f(t)$ be defined on an interval $0 \leq t \leq T$. Then Fourier cosine series with half range expression is of the form [67, Chapter 10]

$$f(t) = c_0 + \sum_{n=1}^{\infty} c_n \cos\left(\frac{n\pi}{T}t\right), \quad 0 \leq t \leq T \tag{B.1}$$

and the coefficients are given by

$$c_0 = \frac{1}{T}\int_0^T f(t)dt \tag{B.2}$$

$$c_n = \frac{2}{T}\int_0^T f(t)\cos\left(\frac{n\pi}{T}t\right)dt\ ,\ n = 1, 2, \ldots. \tag{B.3}$$

For the discrete time signal $f(k)$, $k = 1, \ldots, N$, the sampled version of the continuous time signal $f(t)$ for $0 \leq t \leq T$, the Fourier cosine series with half range expansion of $f(k)$ is defined as

$$f(k) = c_0 + \sum_{n=1}^{N} c_n \cos\left(\frac{n\pi}{N}k\right), \quad k = 1, 2, \ldots, N. \tag{B.4}$$

and the coefficients are expressed as

$$c_0 = \frac{1}{N}\sum_{k=1}^{N} f(k) \tag{B.5}$$

$$c_n = \frac{2}{N}\sum_{k=1}^{N} f(k)\cos\left(\frac{n\pi}{N}\right)\ ,\ n = 1, 2, \ldots, N. \tag{B.6}$$

Appendix C

Representation of a Sinusoid by the Stationary AR Model

Let us consider a real sinusoid $x(n)$ of the form,

$$x(n) = A\cos(2\pi f_0 n + \phi) \tag{C.1}$$

Strictly speaking the use of any time series model for a sinusoid is inappropriate. The AR(2) model is given by

$$x(n) = -a_1 x(n-1) - a_2 x(n-2) + w(n) \tag{C.2}$$

where $w(n)$ is the input white noise with variance σ^2. This model can not generate a sinusoidal time series. A recursive difference equation that can generate $(C.1)$ for $N \geq 0$ is

$$x(n) = 2\cos(2\pi f_0)x(n-1) - x(n-2) + A\cos\phi\delta(n) - A\cos(2\pi f_0 - \phi)\delta(n-1) \tag{C.3}$$

where $x(-2) = x(-1) = 0$ and $\delta(n)$ is the *'dirac delta'* function. The poles of the all pole filter $(C.3)$ are the solution of the equation

$$A(z) = 1 - 2\cos(2\pi f_0)z^{-1} + z^{-2} = 0$$

or,

$$z_p = \exp(\pm j2\pi f_0) \tag{C.4}$$

and hence $(C.3)$ represents an unstable filter.

Inspite of these difficulties, it is still possible to model the auto-correlation function (ACF) of $(C.1)$ as the limiting form ACF of a real AR(2) process. The ACF

of $x(n)$ as given in $(C.1)$ is

$$r_{xx}(k) = \frac{A^2}{2}\cos(2\pi f_0 k) \tag{C.5}$$

This ACF can be approximated by ACF of an AR(2) process with poles at $r\exp(\pm 2\pi f_0)$ if $r \to 1$ $\sigma^2 \to 0$ in such a way that $r_{xx}(0) = \frac{A^2}{2}$.

The ACF of an AR(2) process can be shown as

$$r_{xx}(k) = \sigma^2 \frac{\frac{1+r^2}{1-r^2}\sqrt{1 + \left(\frac{1-r^2}{1+r^2}\right)^2 \cot^2(2\pi f_0)}}{1 - 2r^2\cos(4\pi f_0) + r^4} r^{|k|}\cos(2\pi f_0|k| - \psi) \tag{C.6}$$

where $\psi = \arctan\left(\frac{1-r^2}{1+r^2}\cot(2\pi f_0)\right)$. So, from $(C.6)$ the approximation yields

$$r_{xx}(k) \to \frac{\sigma^2\left(\frac{2}{1-r^2}\right)}{2 - 2\cos(4\pi f_0)}\cos(2\pi f_0 k).$$

From $(C.6)$ and the condition $r_{xx}(0) = \frac{A^2}{2}$ we get,

$$\lim_{r\to 1, \sigma^2\to 0} \frac{\sigma^2}{(1-r^2)(1-\cos(4\pi f_0)} = \frac{A^2}{2} \tag{C.7}$$

In this way a sinusoid may be modelled by an AR(2) process, at least as far the ACF is concerned. For M real sinusoids a real AR(2M) model is required, and for M complex sinusoid complex AR(M) is needed [1, Chapter 5].

Appendix D

A Long Proof

$$
\begin{aligned}
\sum_{i=1+n_0}^{m+n_0} \sin\alpha(i-n_0) &= \sum_{i=1}^{m} \sin\alpha i \\
&= \sum_{i=0}^{m} \sin\alpha i \\
&= \mathrm{Im}\left[\sum_{i=0}^{m} \exp(j\alpha i)\right] \\
&= \mathrm{Im}\left[\frac{\exp(j\alpha(m+1))-1}{\exp(j\alpha)-1}\right] \\
&= \mathrm{Im}\left[\frac{\cos\alpha(m+1)-1+j\sin\alpha(m+1)}{\cos\alpha-1+j\sin\alpha}\right] \\
&= \mathrm{Im}\left[\frac{(\cos\alpha(m+1)-1+j\sin\alpha(m+1))(\cos\alpha-1-j\sin\alpha)}{(\cos\alpha-1)^2+\sin^2\alpha}\right] \\
&= \frac{\cos\alpha\sin\alpha(m+1)-\sin\alpha(m+1)-\cos\alpha(m+1)\sin\alpha+\sin\alpha}{\cos^2\alpha+1-2\cos\alpha-\sin^2\alpha} \\
&= \frac{[sin\alpha(m+1)\cos\alpha-\cos\alpha(m+1)\sin\alpha]+[\sin\alpha-\sin\alpha(m+1)]}{2-2\cos\alpha} \\
&= \frac{\sin\alpha m+2\cos(\alpha+\frac{\alpha m}{2})\cos\frac{\alpha m}{2}}{4\sin^2\frac{\alpha}{2}} \\
&= \frac{\cos\frac{\alpha m}{2}\left[\sin\frac{\alpha m}{2}+\cos(\alpha+\frac{\alpha m}{2})\right]}{2\sin^2\frac{\alpha}{2}}
\end{aligned}
$$

When α is small, the above expression can be reduced to a more simple expres-

sion by substituting summation by integration. The derivation is given below

$$\begin{aligned}
\sum_{i=1}^{m} \sin(\alpha i) &= \sum_{i=0}^{m} \sin(\alpha i) \\
&\simeq \int_{0}^{m} \sin(\alpha t) dt \\
&\simeq \frac{1}{\alpha} [-\cos(\alpha t)]_{0}^{m} \\
&\simeq \frac{1}{\alpha} [1 - \cos(\alpha m)]
\end{aligned}$$

Bibliography

[1] S. M. Kay, *Modern Spectral Estimation.* Englewood Cliffs, N. J.: Prentice Hall, 1988.

[2] Y. Grenier, "Time dependent ARMA modelling of non–stationary signals," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 31, no. 4, pp. 899–911, August 1983.

[3] L. R. Rabiner and R. W. Schafer, *Digital Signal Processing of Speech Signals.* New Jersy: Prentice Hall, 1978.

[4] R. M. Loynes, "On the concept of the spectrum of non–stationary process," *Journal of the Royal Statistical Society, Series B*, vol. 30, no. 1, pp. 1–30, 1968.

[5] S. H. Nawab, "Short–time Fourier transform," in *Advanced Topics in Signal Processing* (J. S. Lim and A. V. Oppenheim, eds.), ch. 6, pp. 289–337, Englewood Cliffs, New Jersy: Prentice Hall, 1 st. ed., 1988.

[6] C. K. Chui, *An Introduction to Wavelets.* Boston: Academic Press, 1992.

[7] J. B. Allen and L. R. Rabiner, "A unified approach to short-time Fourier analysis and synthesis," *Proceedings IEEE*, vol. 65, no. 11, pp. 1558–1564, November 1977.

[8] I. Daubechies, "The wavelet transform, time–frequency localization and signal analysis," *IEEE Trans. on Information Theory*, vol. 36, no. 5, pp. 961–1005, September 1990.

[9] A. Papoulis, *Signal Analysis.* New York: Mc Graw-Hill Book Co., 1977.

[10] T. A. C. M. Classen and W. F. G. Meklenbräuker, "The Wigner distribution – a tool for time–frequency signal analysis – part i : Continuous-time signls," *Philips J. Res.*, vol. 35, no. 3, pp. 217–250, 1980.

[11] T. A. C. M. Classen and W. F. G. Meklenbräuker, "The Wigner distribution – a tool for time–frequency signal analysis – part ii : Discrete-time signls," *Philips J. Res.*, vol. 35, no. 4/5, pp. 276–300, 1980.

[12] T. A. C. M. Classen and W. F. G. Meklenbräuker, "The Wigner distribution – a tool for time–frequency signal analysis – part iii : Relation with other time–frequency transforms," *Philips J. Res.*, vol. 35, no. 6, pp. 372–389, 1980.

[13] L. Cohen, "Time–frequency distribution – a review," *Proceedings IEEE*, vol. 77, no. 7, pp. 941–981, July 1980.

[14] A. J. E. M. Jassen and T. A. C. M. Classen, "On positivity of time–frequency distributions," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 33, no. 4, pp. 1029–1032, August 1985.

[15] A. J. E. M. Jassen, "A note on positivity of time–frequency distributions," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 35, no. 5, pp. 701–703, May 1987.

[16] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. on Signal Processing*, vol. 40, no. 2, pp. 402–412, February 1992.

[17] D. L. Jones and T. W. Parks, "A resolution comparison of several time–frequency presentations," *IEEE Trans. on Signal Processing*, vol. 40, no. 2, pp. 413–420, February 1992.

[18] M. B. Priestley, *Spectral Analysis and Time Series*. London: Academic Press, 1981.

[19] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, N. J.: Prentice Hall, 1988.

[20] R. Kumaresan and D. W. Tuffs, "Estimating the parameters of exponentially damped sinusoids and pole–zero modelling in noise," *IEEE Trans. on Ac-

coustic Speech and Signal Processing, vol. 35, no. 6, pp. 833–840, December 1982.

[21] J. A. Cadzow and M. M. Wu, "Analysis of transient data in noise," *IEE Proceedings, Pt. F*, vol. 134, no. 1, pp. 69–78, February 1987.

[22] S. Mukhopadhyay, *Parameter estimation of complex exponential signals in noise*, Master's thesis, Department of Electrical Engineering, IIT Kanpur, January 1991.

[23] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 34, no. 5, pp. 1081–1089, October 1986.

[24] I. Dologloug and G. Carayannes, "LPC/SVD analysis with zero modelling error," *Signal Processing*, vol. 23, no. 3, pp. 293–298, June 1991.

[25] Y. Grenier, "Rational non–stationary spectra and their estimation," in *1st ASSP Workshop on Spectral Estimation*, (Ontario, Canada), pp. 6.8.1–6.8.8, 1981.

[26] Y. Grenier, "Autoregressive models with time dependent log area ratio," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 36, no. 10, pp. 1062–1612, October 1988.

[27] Y. Grenier, "Non–stationary signal modelling with application to bat echolocation calls," *Accoustica*, vol. 61, no. 3, pp. 155–165, September 1986.

[28] H. M. Teager, "Some observations on oral air flow during phonation," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 28, no. 5, pp. 599–601, October 1980.

[29] G. M. Friesen, *etal* "A comparison of noise sensitivity of nine QRS detection algorithms," *IEEE Trans. on Biomedical Engineering*, vol. 37, no. 1, pp. 85–98, January 1990.

[30] A. C. Guyton, *Text Book of Medical Physiology.* Philadelphia: W. B. Sanders Co., 7th. ed., 1986.

[31] T. S. Rao, "The fitting of non–stationary time-series models with time–dependent parameters," *Journal of Royal Statistics, Series B (Methodological)*, vol. 32, no. 2, pp. 312–322, 1970.

[32] J. A. Cadzow, "Spectral estimation : An overdetermined rational model approach," *Proceedings of IEEE*, vol. 70, no. 9, pp. 907–939, September 1982.

[33] N. C. Huang and J. K. Aggarwal, "Frequency-domain consideration of LSV digital filters," *IEEE Trans. on Circuits and Systems*, vol. 28, no. 4, pp. 279–287, April 1981.

[34] A. H. Gray, "Distance measure for speech processing," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 24, no. 5, pp. 380–391, October 1976.

[35] S. W. A. Gersho and K. Zeger, "Vector quantization techniques in speech coding," in *Advances in Speech Signal Processing* (S. Furui and M. M. Sondhi, eds.), ch. 2, pp. 49–84, New York: Marcel Dekker Inc., 1 st. ed., 1992.

[36] Y. Skiraki and M. Honda, "LPC speech coding based on variable length segment quantization," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 36, no. 9, pp. 1437–1444, September 1988.

[37] S. Wang, "Phonetically–based vector excitation coding of speech at 3.6kbs," in *Proc. IEEE Inter. Conf. on Accoustic Speech and Signal Processing*, (Glasgow), pp. 49–52, 1989.

[38] A. Ralston and P. Rabinowitz, *A First Course in Numerical Analysis*. Singapore: McGraw-Hill, 2 nd ed., 1988.

[39] M. Wan and T. Kailath, "Determination of number of signals by information theoretic criterion," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 33, no. 2, pp. 387–392, April 1985.

[40] L. C. Zhao, P. R. Krishnan and Z. D. Bai, "Determination of number of signals in presence of noise," *Jorunal of Multivariate Analysis*, vol. 20, no. 1, pp. 1–25, October 1986.

[41] J. Rissanen, "Modelling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, September 1978.

[42] N. Jayant, J. Johnston and S. Robert, "Signal compression based on models of human perception," *Proceedings of IEEE*, vol. 81, no. 10, pp. 1358–1422, October 1993.

[43] D. J. Tasell, S. D. Soil, V. M. Kirby and G. P. Widin, "Speech waveform envelop cues for consonants recognition," *Journal of Accoustic Society of America*, vol. 82, no. 4, pp. 1152–1171, October 1987.

[44] J. W. Picone, "Signal modelling techniques in speech recognition," *Proceedings of IEEE*, vol. 81, no. 9, pp. 1215–1247, September 1993.

[45] P. Sircar, *Accurate Parameter Estimation of Damped Sinusoidal Signals Sampled at Non–uniform Spacing.* PhD thesis, Syracuse University, Syracuse, N. Y., December 1987.

[46] N. C. Geçkinli, H. Güngen and M. Etişkol, "Speech synthesis using AM/FM sinusoids and band pass noise," *Signal Processing*, vol. 8, no. 3, pp. 339–361, June 1985.

[47] R. J. McAulay and T. F. Quatier, "Speech analysis/synthesis based on sinusoidal representation," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 34, no. 4, pp. 744–754, August 1986.

[48] J. S. Marques and L. B. Almedia, "Frequency–varying sinusoidal modelling of speech," *IEEE Trans. on Accoustic Speech and Signal Processing*, vol. 37, no. 5, pp. 763–765, May 1989.

[49] N. C. Huang and J. K. Aggarwal, "On linear shift-variant digital filters," *IEEE Trans. on Circuits and Systems*, vol. 27, no. 4, pp. 672–679, August 1980.

[50] H. Tong, *Non–linear Time Series.* Wall Street, Oxford: Oxford University Press, 1990.

[51] E. W. Kamen and K. M. Hafez, "Algebraic theory of linear time–varying systems," *SIAM Journal on Control and Optimization*, vol. 17, no. 4, pp. 500–510, July 1979.

[52] J. F. Kaiser, "On a simple algorithm to calculate the energy of a signal," in *Proc. IEEE Inter. Conf. on Accoustic Speech and Signal Processing*, (Albuquerque, NM), pp. 381–384, 1990.

[53] J. F. Kaiser, "On Teager's energy algorithm and its generalization to continuous signal," in *Proc. 4th IEEE Digital Signal Processing Workshop*, (Mohonk (New Paltz), NY), September 1990.

[54] P. Maragos, J. F. Kaiser and T. F. Quatieri, "On amplitude and frequency demodulation usuing energy operators," *IEEE Trans. on Signal Processing*, vol. 41, no. 4, pp. 1052–1059, April 1993.

[55] P. Maragos, J. F. Kaiser and T. F. Quatieri, "Energy separation in signal modulations with application to speech analysis," Tech. Rep. 91–17, Harvard Robotics Laboratory, Havard University, Cambridge, Massachusetts 02138, November 1991.

[56] A. Potamianos and P. Maragos, "A comparison of the energy operator and Hilbert transform approach to signal and speech demodulation," Tech. Rep. 92–8, Harvard Robotics Laboratory, Havard University, Cambridge, Massachusetts 02138, July 1992.

[57] A. V. Oppenheim and R. W. Schafer, *Discrete–Time Signal Processing*. N.J.: Prentice Hall, 1989.

[58] L. W. Ricketts, J. E . Bridges, J. Miletta, *EMP Radiation and Protective Techniques*. New York: John Wiley, 1976.

[59] A. Swami, "Multiplicative noise models : Parameter estimation using cumulants," *Signal Processing*, vol. 36, no. 3, pp. 355–373, April 1994.

[60] D. Waldo and P. R. Chitrapa, "On Wigner Ville distribution of finite duration signals," *Signal Processing*, vol. 24, no. 2, pp. 231–238, August 1991.

[61] M. Graham, "Latent components in electrocardiogram," *IEEE Trans. on Biomedical Engineering*, vol. 23, no. 3, pp. 220–224, May 1976.

[62] M. E. Wamble, *etal*, "Data compression for storing and transmitting ECG's/VCG's," *Proceedings of IEEE*, vol. 65, no. 5, pp. 702–706, May 1977.

[63] I. S. N. Murthy, *etal*, "Homomorphic analysis of ECG signals," *IEEE Trans. on Biomedical Engineering*, vol. 26, no. 6, pp. 330–344, June 1979.

[64] U. C. Niranjan and I. S. N. Murthy, "ECG component delination by Prony's method," *Signal Processing*, vol. 31, no. 2, pp. 191–202, March 1993.

[65] Sateh M. S. Jalaleddine *etal*, "ECG data compression techniques – a unified approach," *IEEE Trans. on Biomedical Engineering*, vol. 37, no. 4, pp. 329–342, April 1990.

[66] J. McCellan, "Parametric signal modelling," in *Advanced Topics in Signal Processing* (J. S. Lim and A. V. Oppenheim, eds.), ch. 1, pp. 1–57, Englewood Cliffs, New Jersy: Prentice Hall, 1st. ed., 1988.

[67] E. Kreyszig, *Advanced Engineering Mathematics.* New Delhi: Wiley Eastern, 5 th ed., 1983.

Errata

Page No.	Line/Eqn. No.	Written as	To be read as
1	12	advantage	advantages
1	21	non-stationarity	non-stationary
9	5	σ^2	σ_w^2
10	23	spectral distance averaged over time	averaged spectral distance
19	21	for	of
20	24-25	which is	e.g.
21	4	p,q,σ_w^2	p,σ_w^2
22	15	non-stationary	time-dependent
24	1,2,12,16	NSAR	TAR
30	Eqn (3.2)	a_k	$a_k(n)$
39	4	$0.95\angle\frac{\pi}{3}$	$0.95\exp j\frac{\pi}{3}$
47	1	white Gaussian noise	Gaussian white noise
57	5	machinery	machineries
57	6	produces	produce
59	1	different features of the constituent	the constituent
61	3	feature	features
66	21	are	is